Introdução à Estatística

APLICADA À AQUICULTURA

Luan Patrick Moura de Souza

Paulo Roberto Brasil Santos

Suzete Roberta da Silva

Ano 2023

Introdução à Estatística

APLICADA À AQUICULTURA

Luan Patrick Moura de Souza

Paulo Roberto Brasil Santos

Suzete Roberta da Silva

Ano 2023

CIÊNCIAS AGRÁRIAS

Introdução à estatística aplicada à aquicultura

Diagramação: Natália Sandrini de Azevedo
Correção: Yaiddy Paola Martinez
Indexação: Amanda Kelly da Costa Veiga
Revisão: Os autores
Autores: Luan Patrick Moura de Souza
Paulo Roberto Brasil Santos
Suzete Roberta da Silva

Dados Internacionais de Catalogação na Publicação (CIP)

S729 Souza, Luan Patrick Moura de
Introdução à estatística aplicada à aquicultura / Luan Patrick Moura de Souza, Paulo Roberto Brasil Santos, Suzete Roberta da Silva. – Ponta Grossa - PR: Atena, 2023.

Formato: PDF
Requisitos de sistema: Adobe Acrobat Reader
Modo de acesso: World Wide Web
Inclui bibliografia
ISBN 978-65-258-1088-1
DOI: https://doi.org/10.22533/at.ed.881231904

1. Aquicultura. I. Souza, Luan Patrick Moura de. II. Santos, Paulo Roberto Brasil. III. Silva, Suzete Roberta da. IV. Título.

CDD 639.3

Elaborado por Bibliotecária Janaina Ramos – CRB-8/9166

Atena Editora
Ponta Grossa – Paraná – Brasil
Telefone: +55 (42) 3323-5493
www.atenaeditora.com.br
contato@atenaeditora.com.br

DECLARAÇÃO DOS AUTORES

Os autores desta obra: 1. Atestam não possuir qualquer interesse comercial que constitua um conflito de interesses em relação ao artigo científico publicado; 2. Declaram que participaram ativamente da construção dos respectivos manuscritos, preferencialmente na: a) Concepção do estudo, e/ou aquisição de dados, e/ou análise e interpretação de dados; b) Elaboração do artigo ou revisão com vistas a tornar o material intelectualmente relevante; c) Aprovação final do manuscrito para submissão.; 3. Certificam que os artigos científicos publicados estão completamente isentos de dados e/ou resultados fraudulentos; 4. Confirmam a citação e a referência correta de todos os dados e de interpretações de dados de outras pesquisas; 5. Reconhecem terem informado todas as fontes de financiamento recebidas para a consecução da pesquisa; 6. Autorizam a edição da obra, que incluem os registros de ficha catalográfica, ISBN, DOI e demais indexadores, projeto visual e criação de capa, diagramação de miolo, assim como lançamento e divulgação da mesma conforme critérios da Atena Editora.

DECLARAÇÃO DA EDITORA

AGRADECIMENTOS

Primeiramente à Deus pela oportunidade de trabalhar e levar o ensino além dos muros da Universidade.

Aos familiares pelo apoio sempre oferecido, sempre nos incentivando a doar o nosso melhor no trabalho que desenvolvemos na universidade.

À Universidade Federal do Oeste do Pará (UFOPA) que viabiliza, através de editais de financiamento, projetos de pesquisa, ensino, extensão e projetos integrados que, sem seu apoio este material não teria sido viabilizado.

À Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação Tecnológica (Proppit).

À Pró-Reitoria da Cultura, Comunidade e Extensão (PROCCE).

À Pró-Reitoria de Ensino de Graduação (PROEN).

Ao Comitê Gestor de Programas Institucionais (CGPrits) que viabilizaram o edital Edital Nº 03/2019 do Comitê Gestor dos Programas Institucionais (CGPrits) que permitiu a realização do projeto integrado “*Desenvolvimento de alimento alternativo para peixes com resíduos agroindustriais e educação continuada em ciências agrárias em escolas de ensino médio, na UFOPA e produtores rurais*”.

Aos consultores da editora Atena pelos pareceres favoráveis à publicação desde material didático e por suas sugestões valiosas para melhora deste material.

A todos os participantes dos minicursos de Introdução Aplicada à Estatística que nos inspiraram para criação deste material.

Os autores,

APRESENTAÇÃO

É indiscutível que a estatística é essencial em praticamente todas as áreas de conhecimento desde humanas, saúde, engenharias e agrárias como tantas outras. Para as ciências agrárias e, mais especificamente aquicultura, não seria diferente. Este material foi idealizado a partir da observação da dificuldade que estudantes apresentaram em edições de minicursos de **Introdução a estatística aplicada à aquicultura** oferecidos pelo projeto integrado de pesquisa, ensino e extensão "*Desenvolvimento de alimento alternativo para peixes com resíduos agroindustriais e educação continuada em ciências agrárias em escolas de ensino médio, na UFOPA e produtores rurais*". A maioria dos participantes sentiam dificuldades em comandos básicos para aplicar a estatística quer seja no Excel como no programa *Past*. Daí, surgiu a ideia de produzir um material didático de fácil compreensão e que ficasse como um guia para estudantes que estivessem iniciando no mundo incrível da estatística aplicada à aquicultura.

Os autores são professores do curso de Engenharia de aquicultura da Universidade Federal do Oeste do Pará (UFOPA) e um estudante do mesmo curso bolsista do projeto integrado. Destes, o professor Paulo Roberto Brasil Santos tem experiência na área da estatística aplicada às ciências agrárias já tendo sido professor da disciplina Estatística experimental no curso de Engenharia de aquicultura, e com a profa. Suzete Roberta da Silva que possui experiência em projetos de extensão de difusão de conhecimento de aquicultura seguiram com o treinamento do estudante Luan Patrick Moura de Souza que foi muito feliz na sua importante colaboração para que este material existisse.

O objetivo deste material é que ele seja um guia de consulta rápida para estudantes estudando introdução à estatística bem como que estejam iniciando com os estudos estatísticos de seus experimentos acadêmicos.

Suzete Roberta da Silva

Coordenadora do projeto integrado de pesquisa ensino e extensão "*Desenvolvimento de alimento alternativo para peixes com resíduos agroindustriais e educação continuada em ciências agrárias em escolas de ensino médio, na UFOPA e produtores rurais*"

“Não há saber mais ou saber menos: há saberes diferentes”.

Paulo Freire

CAPÍTULO 1

INTRODUÇÃO À ESTATÍSTICA

CONTEXTO HISTÓRICO

A estatística nos dias de hoje, é uma ferramenta indispensável para qualquer profissional que necessita analisar informações. Podemos até pensar que suas técnicas nasceram no mundo contemporâneo atual. Mas a utilização da estatística como suporte para a tomada de decisões é verificada desde do mundo antigo, possivelmente foram os egípcios os primeiros a utlizar a estatística nos registrados pessoas presas em guerras disponíveis para trabalhar nas pirâmides.

Na aquicultura, assim como nas demais ciências agrárias, a estatística é empregada como uma ferramenta poderosa, auxiliando pesquisadores no desenvolvimento de novas técnicas para o aperfeiçoamos da atividade.

O que é Estatística?

"A estatística é uma ciência que se dedica ao desenvolvimento e ao uso de métodos para a coleta, resumo, organização, apresentação e análise de dados"(FARIAS;SOARES;CÉSAR,2003, p. 6).

A estatística se utiliza das teorias probabilísticas para explicar a frequência da ocorrência de eventos, bem como em estudos de observações para estimar ou para possibilitar a previsão de futuros fenômenos. A estatística divide-se em duas partes: descritiva e inferencial. Na área descritiva através de banco de dados, organiza e representa dados em informações, tornando menos complexos e mais fáceis de entender. Já na área inferencial utiliza métodos de estimativas e intervalos de confiança para se observar padrões e estimar futuros fenômenos.

Principais Áreas da Estatística

A Estatística é um ramo da ciência que reúne métodos para o planejamento de experimentos, obtenção, organização, análise, apresentação e interpretação de dados, elaboração de conclusões, ou inferências, sobre a população. Baseia-se nos dados coletados via amostragem, informações pretéritas, bem como em identifica padrões por meio da repetição.

A estatística pode ser dividida basicamente em quatro grandes áreas, conforme a figura 1.1.

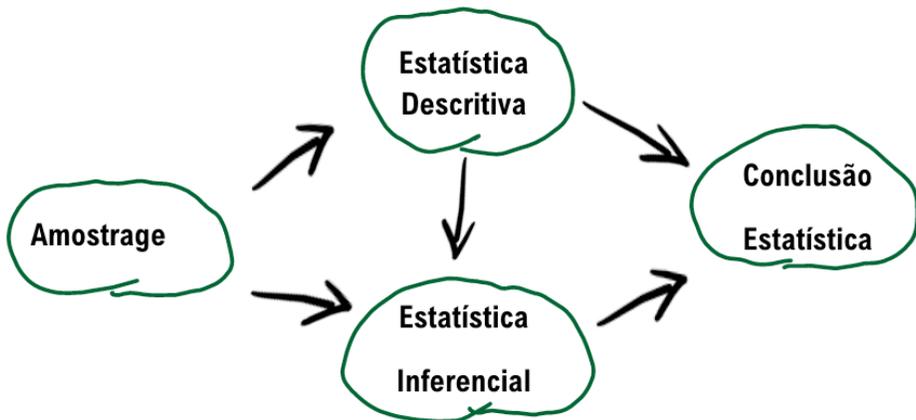



Figura 1.1: Principais áreas da Estatística.

Fonte: Elaborado pelos autores.

Amostragem: É o processo de coletar uma amostra representativa de uma dada população.

Estatística Descritiva: Ao organizar e representar os dados por meio de frequências, medidas de posição e dispersão (Média, Desvio Padrão, Coeficiente de Variação, Erro Padrão), podemos descrever e resumir os dados, para que possamos tirar conclusões a respeito das características da população de interesse.

Estatística Inferencial: Realiza inferências sobre uma população, ou compara duas ou mais populações, por meio de amostras, estima-se intervalos de confiança para responder perguntas hipotéticas, o chamado teste de hipótese.

Conclusão Estatística: Por meio das inferências estatística, deduzimos as propriedades da população, ou comparamos a similaridades entre populações a partir das amostras, para tomar um decisões.

CONCEITOS ESTATÍSTICOS

População

É um conjunto de indivíduos com caraterísticas em comum. O grupo tem características de interesse, para o qual deseja-se fazer uma inferência, para responder uma pergunta científica e tirar uma conclusão.

Exemplos de População: Tilápias *(Oreochromis* spp.*),* de um gênero X;

Todos os tambaquis (*Colossoma macropomum),* do Rio Amazonas.

Em muitos casos a população é inacessível, de difícil acesso. Para contornar a

inacessibilidade da população a estatística utiliza amostras para efetuar as inferências estatísticas. Na figura 1.2 observamos o Conceito de amostra na estatística.

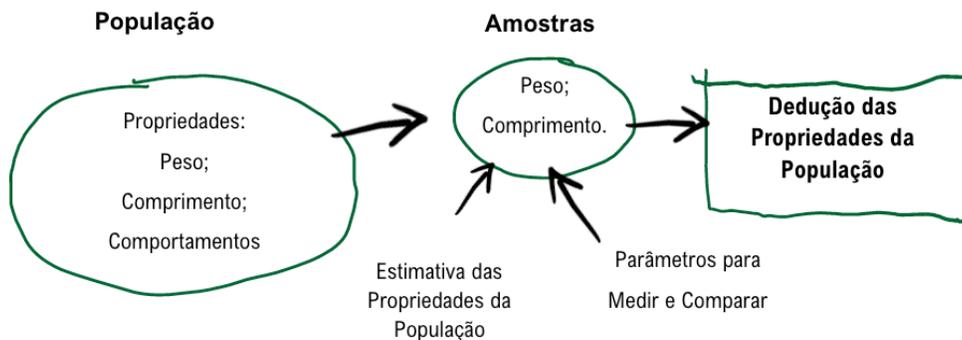



Figura 1.2: Conceito de Amostra na Estatística.

Fonte: Elaborado pelos autores.

Amostra

Uma amostra é uma parte da população. Um subconjunto de integrantes da população, o qual será efetuado a análise estatística, onde permitirá fazer inferências ou tirar conclusões acerca de toda a população. O objetivo de estudar uma amostra da população é diminuir os custos e o tempo da pesquisa. Ao se utilizar amostras adotamos irá ocorrer variações entre os indivíduos da população, denominado erro amostral.

Pesquisa

É um estudo que tenta reunir respostas para responder perguntas hipotéticas, no caso da aquicultura, uma pesquisa irá tentar reunir respostas sobre dúvidas para aperfeiçoar o cultivo, aumentando a lucratividade. O pesquisador é quem elabora a pesquisa, buscando respostas na literatura, ou planeja um experimento afim de tentar responder a dúvida inicial.

Experimento

É o método utilizado para reunir dados com o objetivo de responder as perguntas da pesquisa. Há normas a serem seguidas em um experimento, as chamadas **NBR[1],** estabelece padrões onde o experimento deve se encaixar e/ou seguir.

Conforme a natureza da pesquisa, o experimento muda, o pesquisador pode moldar o experimento a sua realidade. Um pesquisador deve planejar e documentar seu experimento, para que outros pesquisadores possam repetir o experimento e obter os mesmos resultados do experimento inicial.

Tratamento

São os diferentes procedimentos ou interferências, aplicados em um grupo de estudo, podendo ou não gerar efeitos, os quais são mensurados através de variáveis. Em alguns experimentos, há grupos não recebem procedimentos ou qualquer interferência, chamados grupo controle, serve como elemento de comparação com o grupo de estudo.

O alvo do experimento é medir e comparar os efeitos dos diferentes tratamentos nas amostras. Por exemplo: Em uma pesquisa sobre nutrição de tambaqui (*Colossoma macropomum*), o experimento foi desenvolvido para comparar o ganho de peso, entre uma ração comercial e uma ração artesanal, os tratamentos serão cada tipo de ração, como são 2 tipos de rações (comercial e artesanal), teremos um experimento com 2 tratamentos, 2 amostras de tambaqui receberão diferentes rações.

Parcela

É a menor unidade de um experimento, é a unidade onde é aplicado o tratamento. Na Aquicultura, as parcelas variam dependendo da pesquisa. Pode ir de tudo de ensaio, a um tanque de cultivo, ou até uma fazenda com dezenas de tanques de cultivo, com milhares ou milhões de animais cultivados. Uma parcela também pode ser um único animal infectado por parasitas.

Em um experimento a análise estatística requer repetição para observar padrões que repetem e estimar futuro fenômenos. São as parcelas que vão proporcionar a repetição.

Exemplo: Um experimento com 2 tratamentos (ração comercial e ração artesanal), irá ter 4 parcelas, para que os tratamentos se repitam pelos menos uma vez no experimento.

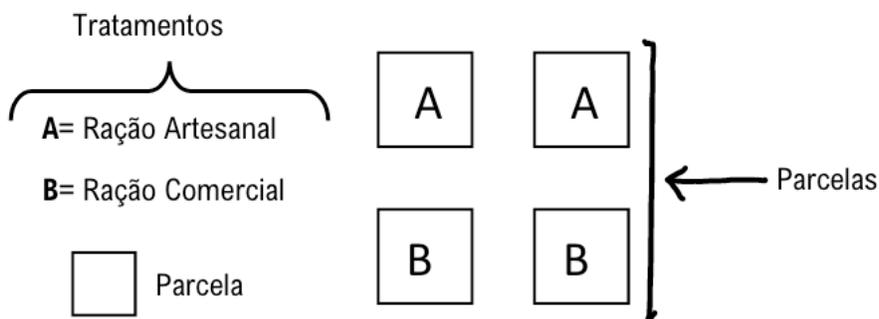



Figura 1.3: Disposição dos Tratamentos nas Parcelas.

Delineamento Experimental

É uma maneira de organizar o experimento, um plano de como dispor os diferentes tratamentos nas parcelas do experimento. Através da casualização dos tratamentos, as parcelas são arranjadas para haja repetições. Tudo de forma aleatória, para que possamos estimar a interferência do ambiente no experimento.

Entre os tipos de delineamento estão: delineamento inteiramente casualizado, o delineamento em blocos casualizados, o delineamento em quadrado latino etc.

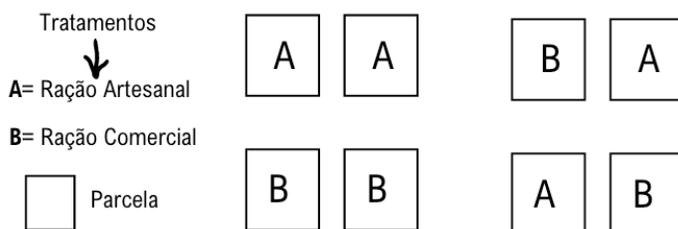



Figura 1.4: Formas de dispor os Tratamentos nas Parcelas.

Casualização

É a distribuição de forma aleatória dos tratamentos nas parcelas por meio de sorteio, evitando que um tratamento seja favorecido ou prejudicado por fatores externos, como os efeitos ambientais.

VARIÁVEIS

Referem-se às informações, características, atributos, índices, passíveis de serem mensurados e que serão utilizados para descrever ou comparar.

Na aquicultura são utilizadas variáveis para mensurar índices zootécnicos dos animais cultivados como, crescimento, parâmetros hematológicos, presença de infestação de algum parasita. Há parâmetros químicos da água como, fósforo, nitrogênio, oxigênio dissolvido, pH, alcalinidade, entre outros.

Classificação de variáveis

As variáveis podem ser classificadas, conforme sua escala de medida.

Qualitativas: *as variáveis não são mensuradas numericamente, mas classificadas em categorias ou grupos, que são denominados de níveis.*

- Nominais: são compostas por nomes, grupos, rótulos, ou categorias sem uma

ordem definida.

- Ordinais: os nomes, grupos, rótulos, ou categorias possuem uma ordem definida.

Quantitativa: as variáveis são mensuradas numericamente, se dividir em:

- Discreta: os valores obtidos são números inteiros. Normalmente são resultados de contagens e não possuem casas decimais.
- Contínua: a medida pode apresentar casas decimais, pois sempre é possível obter um valor intermediário entre outros dois valores quaisquer.

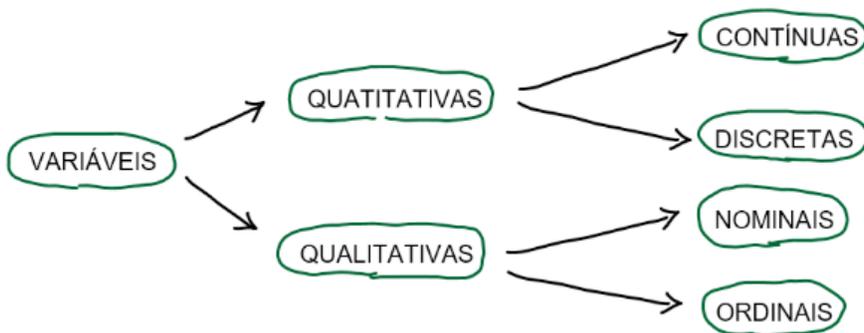



Figura 1.5: Classificação de variáveis.

ERROS

Em experimento estatísticos, os erros são a diferença entre valor observado e o valor real, como não se sabe o valor real, não sabemos o tamanho do erro, apenas estima-se o tamanho deste erro. Assume-se que erros sempre irão acontecer no experimento, pois não é possível eliminá-los por completo. Por isso, adota-se medidas de precisão e intervalos de confiança para estimar o tamanho dos erros.

Erros Aleatórios

São os erros onde cada resultado sofre interferência em torno do valor real, não podem ser eliminados, apenas pode ser estimado. Por não se conhecer o valor real, apenas estima-se o tamanho do erro. E o tipo de erro que ocorre o quando se mede variáveis, pois o ato de mensurar envolve erros do equipamento não ser muito preciso, ou do operador por influenciar no procedimento.

Exemplo: Na aquicultura acontece por exemplo, quando ao pesar de um animal, que não fica parado na balança, ocasionando uma variação do peso real, obrigando o operador

a escolher um valor fixo dentro da variação que ocorre.

Erros Sistemáticos

São erros que deslocam os resultados de forma semelhante, se distanciando do valor real, gerando tendências, falsos resultados no experimento. Esse tipo de erro ocorre, por má seleção das amostras, por equipamentos danificados ou mal calibrados.

Os erros sistemáticos podem ser identificados e eliminados do experimento, mas não podem ser estimados. Quando ocorrem em um experimento, devem ser identificados e solucionados, e como não se sabe o tamanho do deslocamento dos resultados, não se utiliza os resultados nas análises estatísticas. O experimento deve ser refeito, gerando novos resultados para realizar as análises.

Erro Amostral

É a diferença entre o resultado da amostra e o valor real da população inteira. Ao selecionar aleatoriamente indivíduos de uma mesma população, surgirão indivíduos semelhantes, mas não 100% iguais cada um terá características próprias.

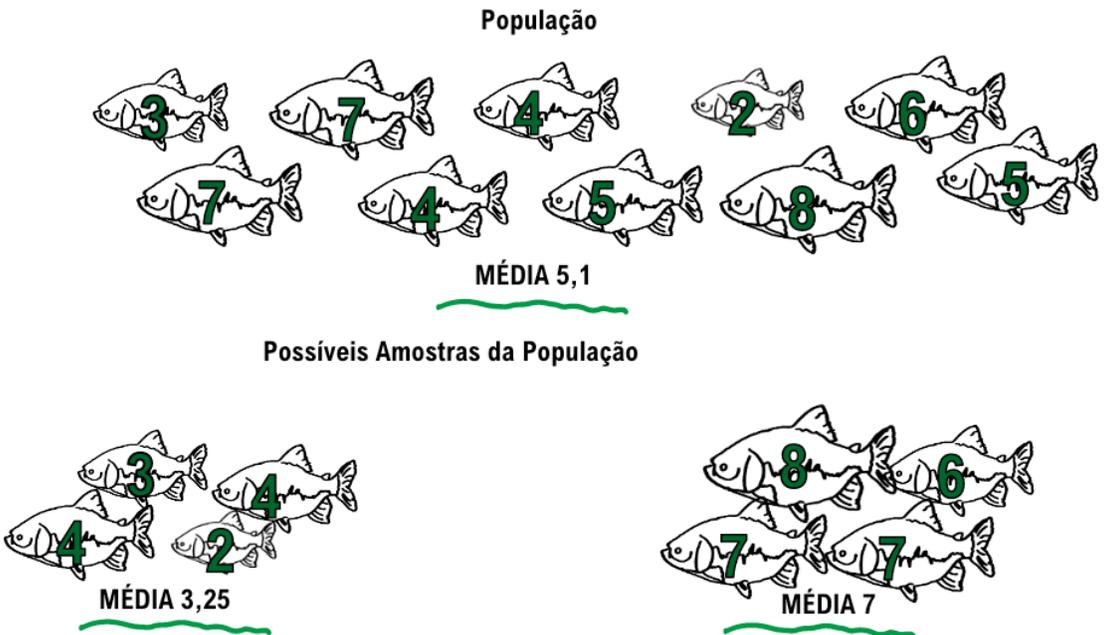



Figura 1.6: Possíveis médias da População

INFERÊNCIA

Métodos de fazer afirmações a partir de dados coletados de amostras a respeito da

população sobre uma ou mais caraterísticas. Verificamos uma dada hipótese sobre um ou mais populações que queremos tirar inferências, ou comparar.

Intervalo de confiança (IC)

A probabilidade de os resultados de um parâmetro da amostra estar próximo ao parâmetro da população inteira, é determinada pelo intervalo de confiança (IC), é um intervalo de valores obtidos da amostra, que têm a probabilidade de conter o valor real de um parâmetro da população inteira. Expresso em porcentagem, exemplo: um intervalo de confiança de 95%, indica que 95% das amostras contém o verdadeiro valor do parâmetro da população.

Teste de Hipótese

É uma ferramenta estatística para tomar uma decisão, de aceitar ou rejeitar uma hipótese através dos resultados obtidos das amostras. Há duas hipóteses no teste de hipótese, a hipótese nula (H_0), afirma uma igualdade, e a hipótese alternativa (H_1), afirma a diferença. No teste de hipótese aceita-se ou rejeita-se H_0 , se aceitar H_0 , assume-se que dados não tem diferença significativa.

Na tomada de uma decisão, entre aceitar ou rejeitar H_0 , pode acontecer um erro. O chamado **ERRO DO TIPO I**, ocorrem quando se rejeita a hipótese nula (H_0), quando na verdade ela é verdadeira, isso significa dizer que assume-se que há uma diferença significativa entre os resultados, quando na verdade não tem.

Hipótese:
Utilizar Ração Artesanal na alimentação de Tambaquis tem o mesmo desempenho da Ração Comercial, diminuindo os custos com Arraçoamento.

Teste de Hipótese

(1.1):

$$H_0 \Rightarrow A = B$$

$$H_1 \Rightarrow A \neq B$$

A= Ração Artesanal

B= Ração Comercial

Exemplo 1.1: Teste de Hipótese para comparar duas rações.

Ao aceitar assume-se que as duas rações (Ração Artesanal e Comercial), tem desempenho semelhante.

Nível de Significância (α)

Utilizado nos testes estatísticos, é a probabilidade de rejeitar a hipótese nula quando de fato é verdadeira, é um índice probabilístico que impede de cometer o **ERRO DO TIPO I**. É o pesquisador que decide o nível de significância que irá adotar no experimento, em geral é utilizando α = 0,05 ou 0,01, isso significa que o pesquisador tem 5% ou 1% de chance de cometer o Erro tipo I.

O nível de significância é o complemento do intervalo de confiança. Isto é, se o nível de significância for 0,05 (5%) o intervalo de confiança de 95%, o pesquisar tem 95% de certeza no experimento.

p-Valor

É um valor probabilístico, utilizado na maioria dos nos testes estatísticos, onde se verifica em um teste de hipótese a probabilidade, de obter um valor entro da região crítica do teste, rejeitando (), ou fora da região crítica, aceitando (), conforme o nível de significância (α) adotado no teste.

Em geral se o **p-Valor** for **≥** que o nível de significância (**α = 0,05**), aceita-se , pois o p-Valor estar fora a região crítica. Se o **p-Valor** for **≤** que o nível de significância (**α**), rejeita-se , pois o p-valor estando dentro da região crítica, demonstra que há diferença significativa entre os grupos comparados.

Ao utilizar um nível de significância de 0,05 (5%), em teste estatístico, assume-se que um resultado de p-Valor de (**≤** 0,05), há diferença estatística.

CAPÍTULO 2

PARÂMETROS ESTATÍSTICOS

A estatística reúne parâmetros para a coleta de dados (variáveis), tornando-os compreensíveis. Neste capítulo veremos as possíveis formas de calcular parâmetros estatísticos coletados da amostra – como a média, proporção e desvio padrão.

MÉDIA

A média anivela todos os indivíduos, para que todos apresentem o mesmo valor da característica mensurada, é dado por:

(2.1): $$\mu = \frac{\sum X}{N}$$ Fórmula da Média População

Onde μ é a média, $\sum X$ é a somatória da característica da população e N é o tamanho da população.

Com a finalidade de diminuir os custos e o tempo da pesquisa, utiliza-se uma amostra da população. No caso da a fórmula para calcular a média da amostra, definida por:

(2.2): $$\bar{X} = \frac{\sum X}{n}$$ Fórmula da Média Amostra

Onde $\sum X$ é a soma dos valores de X na amostra e n é o tamanho da amostra.

EXEMPLO 2.1: Abaixo na **Figura 2.1,** temos uma amostra de peso de 7 juvenis de tambaqui em gramas.

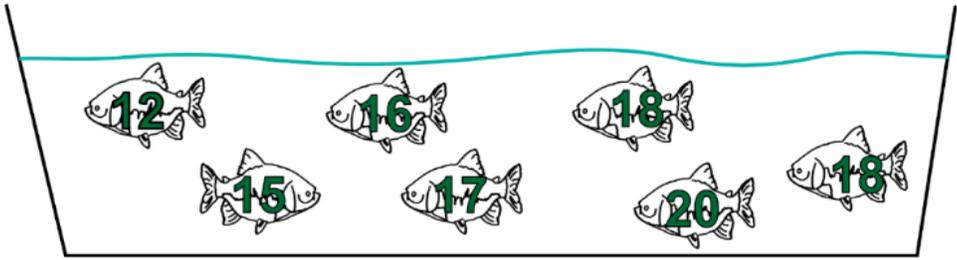



Figura 2.1: Amostra de Pesos de Tambaqui em gramas.

(2.4)
$$\bar{X} = \frac{\sum X}{n} = \frac{(12+15+16+17+18+18+20)}{7}$$
$$\bar{X} = \frac{116}{7}$$
$$\bar{X} = 16{,}6 \text{ g}$$

DESVIO PADRÃO

O desvio padrão é uma medida de dispersão que mede o quanto os valores variam dos valores centrais da massa de dados. Se desvio padrão for pequeno, significa que os valores estão próximos da média, enquanto se o desvio padrão for grande, significa que os valores estão longe da média.

Quanto menor for o desvio padrão mais homogênea é a amostra, na aquicultura é mais vantajoso ter um cultivo de animais mais homogêneos, pois são mais valorizados no comércio.

O desvio padrão de uma população é definido como:

(2.3):
$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}}$$
:Fórmula Desvio Padrão População

O desvio padrão de uma amostra é:

(2.4):
$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{n - 1}}$$
:Fórmula Desvio Padrão Amostra

EXEMPLO 2.2: Para calcular o desvio padrão da amostra de pesos da **Figura 2.1** foram organizamos na tabela (Tabela 2.1):

n	$(X = \bar{X})^2$
12	20,89796
15	2,46939
16	0,32653
17	0,18367
18	2,04082
18	2,04082
20	11,75510

Na primeira coluna, estão os dados. Na segunda coluna, subtraímos a média 16,6, obtida no exemplo 2.1, elevamos o resultado ao quadrado. A última linha é a soma coluna.

Ao realizar o cálculo do desvio padrão com as amostras da **Tabela 2.1**, iremos obter a seguinte expressão:

(2.5):

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{n-1}} = \sqrt{\frac{39{,}71429}{7-1}} = \sqrt{6{,}6190} = 2{,}572$$

Portanto, o desvio padrão desta amostra é:

$$S = 2{,}57g$$

O desvio padrão possui a mesma unidade dos dados.

Assim, a média da amostra de pesos do exemplo 1.1, é 16,6g e o desvio padrão é 2,6 g. Isto significa que os dados estão mais ou menos na faixa de 16,6 ± 2,6g, ou seja, de 14g a 19,2g.

VARIÂNCIA

A variância é uma medida de dispersão que mostra a distância entre os valores entorno da média. Quanto menor é seu valor menor será a variação dos valores amostrados em relação à média. Dados com variância baixa são mais homogêneos.

(2.6):

$$s^2 = \frac{\sum(X - \bar{X})^2}{n-1}$$:Fórmula Variância Amostra

Uma forma mais simples de entender a fórmula da variância é:

(2.7):

$$s^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + (X_3 - \bar{X})^2 + \ldots + (X_n - \bar{X})^2}{n - 1}$$

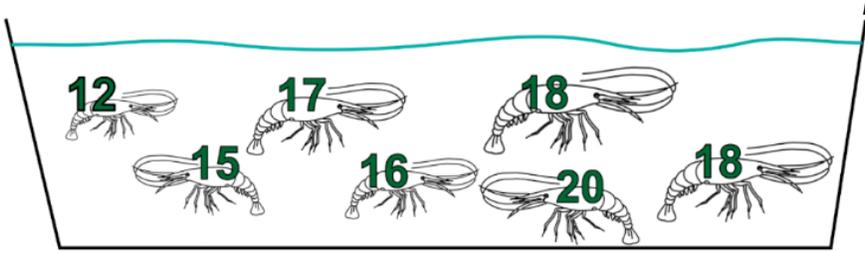



Figura 2.2: Amostra de Pesos de Camarão em gramas.

EXEMPLO 2.3: Para calcular a variância para os dados de indivíduos da **Figura 2.2**, sabe-se a média é 16,6, teremos algo como:

$$s^2 = \frac{(12-16{,}6)^2+(15-16{,}6)^2+(16-16{,}6)^2+}{7-1} \Rightarrow$$

$$\Rightarrow \frac{(17-16{,}6)^2+(18-16{,}6)^2+(18-16{,}6)^2+(20-16{,}6)^2}{7-1}$$

(2.8):

$$s^2 = \frac{(-4{,}6)^2+(-1{,}6)^2+(-0{,}6)^2+(0{,}4)^2+(1{,}4)^2+(1{,}4)^2+(3{,}4)^2}{6}$$

$$s^2 = \frac{21{,}16+2{,}56+0{,}36+0{,}16+1{,}96+1{,}96+11{,}56}{6}$$

$$s^2 = \frac{39{,}72}{6}$$

$$s^2 = 6{,}62\ g^2$$

Após o cálculo da variância, a unidade medida dos dados se eleva ao quadrado (x^2), como a **Figura 2.2** é peso em gramas(g) o resultado do **Exemplo 2.3** será g^2.

COEFICIENTE DE VARIAÇÃO

É a razão entre desvio padrão e média. É uma forma de expressar a variação dos dados em termos percentuais.

A fórmula do coeficiente de variação é dada por:

(2.9):

$$CV = \frac{s}{\bar{X}} \times 100$$

:Fórmula Coeficiente de Variação

Onde s é o desvio padrão, $\bar{X}$ é a média, o resultado da divisão é multiplicado por 100 para obter um percentual (%).

EXEMPLO 2.4: Para os dados de peso da **Figura 2.2**, temos um desvio padrão s = 2,6 g com uma média amostral de 16,6 g. Então cálculo do coeficiente de variação dos pesos é:

(2.10):
$$CV = \frac{2,6}{16,6} \times 100 = 15,6\%$$

ERRO PADRÃO

É uma medida que estima a credibilidade da média, uma estimativa do intervalo de confiança da média.

A fórmula do Erro padrão é dado por:

(2.11):
$$EP = \frac{s}{\sqrt{n}}$$

Onde s é o desvio padrão, n é o tamanho da amostra.

Exemplo 2.5: Como os dados da **Figura 2.2**, a média é 2,6 g e o tamanho da amostra é 7, então teremos o seguinte cálculo para o erro padrão.

(2.12):
$$EP = \frac{2,6}{\sqrt{7}} \qquad EP = 0,98$$

AMPLITUDE

É a diferença entre o maior valor e o menor valor da amostra.

A fórmula da Amplitude (R) dado por:

(2.13):
$$R = \text{MÁXIMO} - \text{MÍNIMO}.$$

CAPÍTULO 3

MANIPULAÇÃO DE DADOS NO EXCEL

APRESENTAÇÃO DO MICROSOFT EXCEL

O pacote computacional Excel é um pacote de uso geral para cálculos na administração, contabilidade, ciência e outras áreas. Seu uso é bastante simples e disseminado.

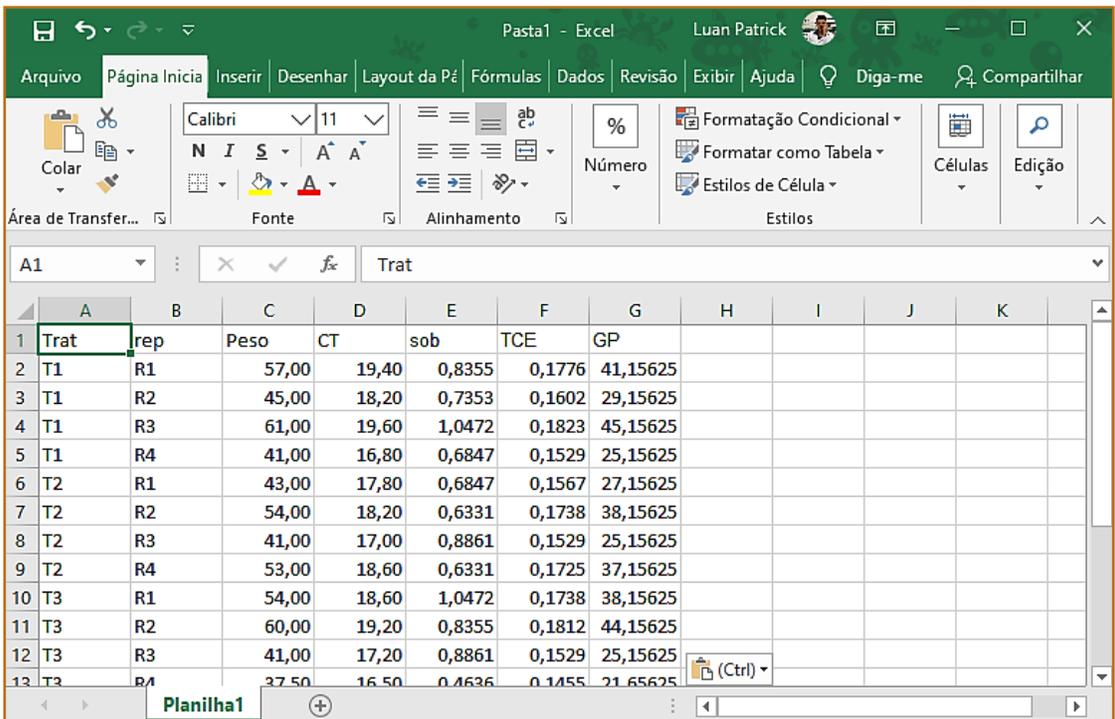
O Excel trabalha com o sistema de planilhas. Uma planilha é formada de células, cada uma delas identificada por uma letra (correspondente à coluna) e um número (correspondente à linha). Cada célula pode conter números, texto ou fórmulas. A **Figura 3.1** mostra uma planilha do Excel.

	A	B	C	D	E	F	G
1	Trat	rep	Peso	CT	sob	TCE	GP
2	T1	R1	57,00	19,40	0,8355	0,1776	41,15625
3	T1	R2	45,00	18,20	0,7353	0,1602	29,15625
4	T1	R3	61,00	19,60	1,0472	0,1823	45,15625
5	T1	R4	41,00	16,80	0,6847	0,1529	25,15625
6	T2	R1	43,00	17,80	0,6847	0,1567	27,15625
7	T2	R2	54,00	18,20	0,6331	0,1738	38,15625
8	T2	R3	41,00	17,00	0,8861	0,1529	25,15625
9	T2	R4	53,00	18,60	0,6331	0,1725	37,15625
10	T3	R1	54,00	18,60	1,0472	0,1738	38,15625
11	T3	R2	60,00	19,20	0,8355	0,1812	44,15625
12	T3	R3	41,00	17,20	0,8861	0,1529	25,15625
13	T3	R4	37,50	16,50	0,4636	0,1455	21,65625

Figura 3.1: Área de trabalho do Excel

Área de Trabalho do Excel

Ao ser carregado, o Excel exibe sua tela de trabalho mostrando uma planilha em branco com o nome de Pasta 1. A área de trabalho do EXCEL 2019 é composta por diversos elementos, entre os quais podemos destacar os seguintes. **Figura 3.1** mostra os elementos do Excel.

Células

Uma planilha é composta por células. Uma célula é o cruzamento de uma coluna com uma linha. A função de uma célula é armazenar informações que podem ser um texto, um número ou uma fórmula que faça menção ao conteúdo de outras células. Cada célula é identificada por um endereço que é composto pela letra da coluna e pelo número da linha. A **Figura 3.2**, mostra uma célula selecionada.

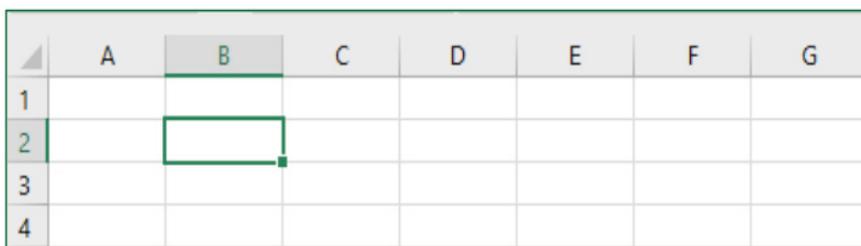



Figura 3.2: Exemplo de Célula.

Guia das Planilhas

Na parte inferior esquerda do Excel, encontra-se as guias das planilhas, servem para organizar as planilhas quando um arquivo o Excel tem mais de uma planilha. Ao aprestar no botão + , são criadas novas planilhas com os nomes Plan1, Plan2, etc., mas podem ser renomeados. A **Figura 3.3** mostra as guias de uma planilha.

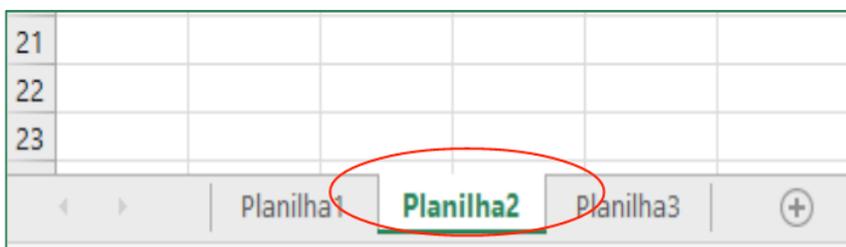



Figura 3.3: Guias das Planilhas. Planilha selecionada destacada circulada em vermelho.

Barra de fórmulas

Tem como finalidade exibir o conteúdo da célula atual, permite ainda a edição do conteúdo de uma célula nela podemos inserir fórmulas para realização de cálculos. A **Figura 3.4** mostra a barra de fórmulas.

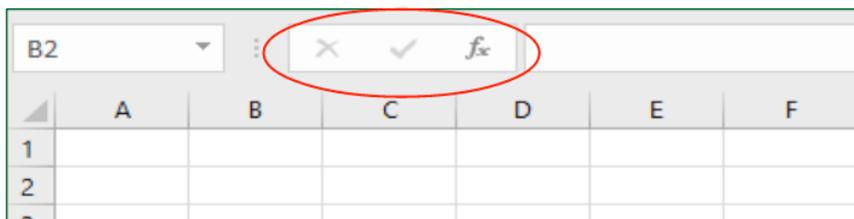



Figura 3.4: Barra de fórmulas. Está destacado circulado em vermelho.

Auto Preenchimento

Tem como finalidade de preencher células com dados quem sigam um padrão ou que tenham base em dados de outras células.

Em uma ou mais célula que deseja usar como base para preenchimento de outras células.

Para uma série como 1, 2, 3, 4, 5...., digite 1 e 2 nas primeiras duas células. Para a série 2, 4, 6, 8..., digite 2 e 4 nas primeiras duas células. Para a série 2, 2, 2, 2..., digite 2 na primeira célula.

Selecione as primeiras duas células da série, exemplo 1, 2 ou 2, 4. Arraste a alça de preenchimento até a célula que deseja inserir o autopreenchimento.

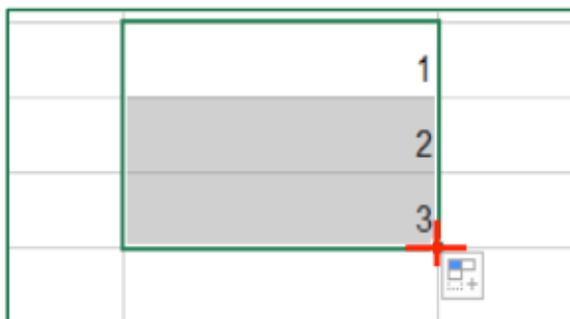



Figura 3.5: Auto Preenchimento em série (1, 2, 3 ...). Em vermelho indica o local onde deve-se clicar para arrastar e obter o auto preenchimento em série.

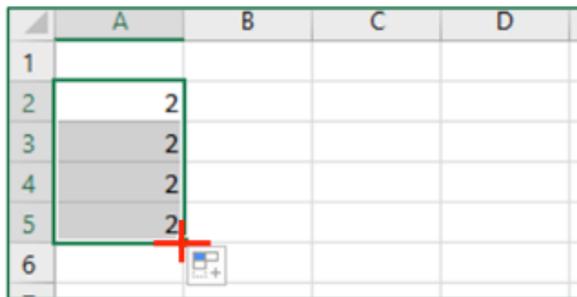



Figura 3.6: Auto Preenchimento em série 2, 2, 2. Em vermelho indica o local onde deve-se clicar para arrastar e obter o auto preenchimento em série.

A Barra de Títulos

É onde mostra o nome do arquivo (aqui, «Pasta1», é o título padrão. Você será capaz de dar um novo nome a planilha quando salvar. Há opções no canto superior esquerdo, onde podemos personalizar, mas por padrão temos: Salvar, Desfazer, Refazer e o botão para personalizar a barra de títulos.

Figura 3.7: Mostra a Barra de Títulos

Ao lado dos botões de minimizar, maximizar e fechar, no canto superior direito, nas versões novas do Excel tem como conectar uma conta e-mail, possibilitando trabalhar documentos em nuvem.

O sistema de menu em Guias

O sistema de menus em guias organiza e facilita a navegação no Excel, ao clicar em uma guia, abre vários outros comandos, denominados Faixas de Opções. Cada guia tem suas faixas de opções. Na guia Página Inicial terá as principais opções, as demais guias terão comando específicos para cada funcionalidade que o Excel pode fazer.

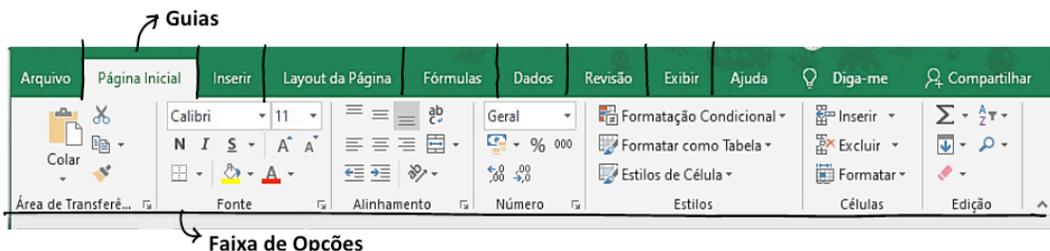



Figura 3.8: Mostra a Faixa de Opções.

3.1.8 Menu Arquivo

Arquivo No Excel 2019, há uma guia chamada "Arquivo" no canto superior esquerdo. Ao clicar na guia aparece um menu. A partir deste menu abre outras funcionalidades essenciais **Figura 3.9**, tais como: Criar uma nova planilha, abrir um arquivo existente, salvar a planilha, imprimir, exportar a planilha. Há também opções da conta e opções de ajustes do Excel.

Informações
Planilha
D: » ebook
Carregar | Compartilhar | Copiar caminho | Abrir Local do Arquivo

Página Inicial
Novo
Abrir
Informações
Salvar
Salvar como
Histórico
Imprimir
Compartilhar
Exportar
Fechar
Conta
Comentários
Opções

Proteger Pasta de Trabalho
Controle que tipos de alterações as pessoas podem fazer nesta pasta de trabalho.

Inspecionar Pasta de Trabalho
Antes de publicar este arquivo, saiba que ele contém:
- Propriedades do documento, caminho da impressora, nome do autor, datas relacionadas, dados de imagem recortados e caminho absoluto
- Conteúdo que pessoas com deficiência não conseguirão ler

Gerenciar Pasta de Trabalho
Não existem alterações não salvas.

Opções de Exibição do Navegador
Controle o que os usuários podem ver quando esta pasta de trabalho é exibida na Web.

Propriedades
Tamanho 585KB
Título Adicionar um título
Marcas Adicionar marca
Categorias Adicionar categoria
Datas Relacionadas
Última Modificação Hoje, 17:51
Criada em 17/11/2019 19:53
Última Impressão
Pessoas Relacionadas
Autor Paulo Brasil
Adicionar um autor
Última Modificação por Luan Patrick
Documentos Relacionados
Abrir Local do Arquivo

Figura 3.9: Mostra as opções menu Arquivo

Barra de Acesso Rápido

No lado superior esquerdo, encontramos a barra de acesso rápido vários, é um

espaço para acrescentar atalhos de configuráveis, para agilizar o fluxo de trabalho. Além executar tarefas comuns, como salvar e desfazer, sem ter que encontrá-los em um menu, há como acrescentar dezenas de outros comandos.

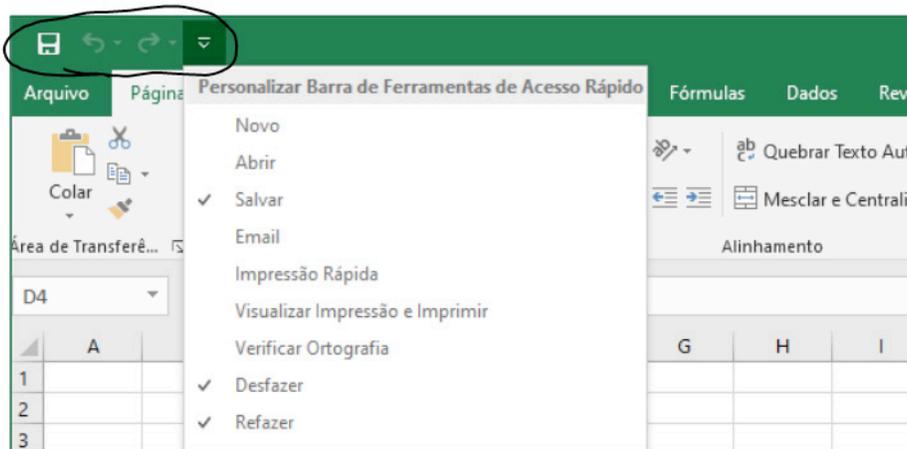



Figura 3.10: Barra de Acesso Rápido.

Página Inicial

Os comandos mais utilizados no Excel são também os mais acessíveis. Alguns destes comandos estão disponíveis na aba Página Inicial. Na **Figura 3.11** mostra alguns desse comandos.

A guia Página Inicial oferece opções que podem mudar o tipo de letra, tamanho, cor, alinhamento, organização e estilo do texto na planilha e células individuais.

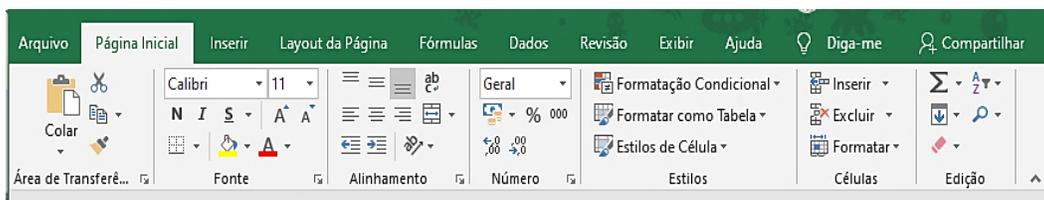



Figura 3.11: Guia Página Inicial.

Cada guia é subdividida em categorias de comandos. Como observando na **Figuro 3.11**, pode-se encontrar as categorias de comandos: *Área de Transferência, Fonte, Alinhamento, Número, Estilo, Células e Edição.* Cada uma destas categorias possui comandos específicos.

EQUAÇÕES NO EXCEL

O Excel pode somar, subtrair, multiplicar, dividir, encontrar a média, e executar funções de contagem gerais sobre os dados numéricos que forem selecionados." Para a estatística o Excel é uma ferramenta importante principalmente para organização dos dados e criação de gráficos, algumas das principais funções e comandos matemáticos no Excel são:

- SOMA E SUBTRAÇÃO
- MULTIPLICAÇÃO E DIVISÃO
- MÉDIA
- MAIOR VALOR (MÁXIMO)
- MENOR VALOR (MÍNIMO)
- CONTAGEM DE NÚMEROS

Para iniciar uma função ou uma simples fórmula matemática é sempre necessário utilizar sinal de igual (**=**), e o Excel irá identificar a fórmula ou operação matemática a ser utilizando.

Soma

Basicamente, é possível realizar a soma de duas maneiras, utilizando o sinal de mais (**+**) ou a função *=SOMA.* Normalmente, usa-se a função *=SOMA* para somar intervalos contínuos ou não de células, e a soma com o sinal de mais (+) para somar entre poucas células.

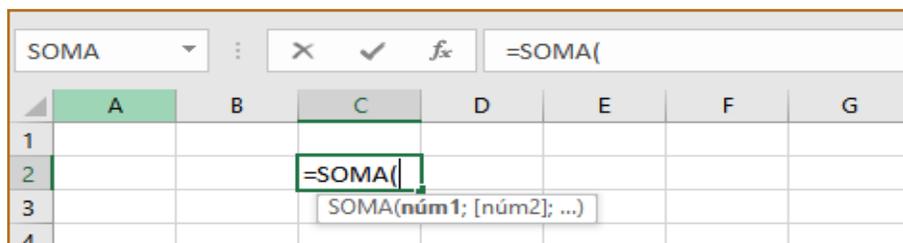



Figura 3.12: Função Soma.

Para realizar o cálculo de soma sinal de mais (+) é necessário adicionar o sinal de mais (+) entre cada elemento da soma. Se a soma for entre os endereços das células aonde estão os valores, a soma irá assumir algo parecido com a **Figura 3.13**.

Figura 3.13: Soma

Exemplos:

=C1+C2 (soma o conteúdo das células C1 e C2);

=C1+C2+C3+C3+... (soma o conteúdo das células C1, C2, C3 e C3).

Há também com utilizar a função *=SOMA*. Na **Figura 3.14** e **Figura 3.15** são exemplos diferentes de como utiliza a função *=SOMA* do Excel.

Figura 3.14: Função =SOMA de células

=SOMA(A1 **;** A2), com ponto e vírgula(;), soma somete o conteúdo da célula A1 com A2.

Figura 3.15: Função =SOMA com intervalos

=SOMA(A1 **:** A2), com dois pontos (:), soma o conteúdo das células que estão no intervalo de A1 até A2.

Figura 3.16: Função *=SOMA* com intervalos e células isoladas

Com a função =*SOMA* é possível somar intervalos (:) com outros intervalos e/ou somar células isoladas utilizando (;) para separar.

Subtração

Não Há uma função *"=SUBTRAIR"* no Excel, para fazer uma subtração simples, utilizamos sinal de subtração (-). Use a função SOMA e converta quaisquer números que você deseja subtrair em valores negativos.

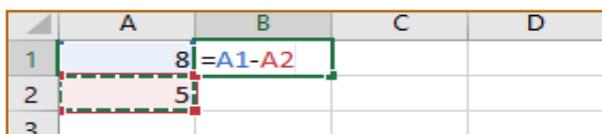



Figura 3.17: Subtração

Exemplos:

= A1 – A2 (subtrai o conteúdo das células A1 e A2).

=A1-A2-D3-E3 (subtrai o conteúdo das células A1, A2, D3 e E3).

Multiplicação

Para realizar multiplicações no Excel há duas maneiras, utilizando a função *=MULT* multiplica o intervalo das células selecionadas ou utilizando o símbolo asterisco (*), por exemplo **=**A1 * A2.

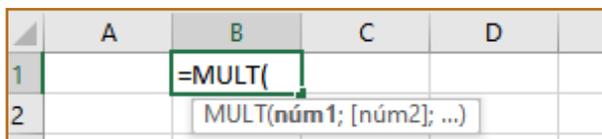



Figura 3.18: Multiplicação

Exemplos:

=C1*C2 (multiplica o conteúdo das células C1 e C2);

=C1*C2*C3*C3 (multiplica o conteúdo das células C1, C2, C3 e C3).

Figura 3.19: Função *=MULT*

=MULT(A1;A2) com ponto e vírgula (;), multiplica somente o conteúdo da célula A1 com A2.

	A	B	C	D
1	8	=MULT(A1:A4)		
2	5	MULT(núm1; [núm2]; ...)		
3	4			
4	3			

Figura 3.20: Função *=MULT* com intervalos

=MULT(A1:A4) com dois pontos (:), multiplica o conteúdo das células que estão no intervalo de A1 até A4.

Com a função *=MULT* é possível adicionar valores individuais, referências de célula ou intervalos, ou uma mistura dos três ao mesmo tempo.

Figura 3.21: Função =MULT composta

Divisão

Para realizar uma subtração no Excel não há uma função "*=DIVIDIR*", para executar essa tarefa utiliza-se o símbolo barra (/). Por exemplo, se você digitar **= 10/5** em uma célula, a célula exibirá **2**.

Figura 3.22: Divisão

Exemplo:

=A1/A2 (divide o conteúdo das células A1 pelo conteúdo da A2).

Média

Para calcular a média de um intervalo de células utiliza-se a função *=MÉDIA*. Está função se aplica em intervalos de células. Por exemplo a função ***=****MÉDIA* (A1:A5) calculará a média das células do intervalo A1 a A5.

Figura 3.23: Função *=MÉDIA*

Exemplo:

=MÉDIA(A1:A3) (calcula a média dos valores das células do intervalo de A1 até A3).

Figura 3.24: Exemplo unção *=MÉDIA* cálculo de intervalo

3.2.6 Máximo e Mínimo

A função *=MÁXIMO* é utilizada para aponta qual é o maior valor dos intervalos das células.

Figura 3.25: Função *=MÁXIMO*

A função *=MÍNIMO* tem a mesma finalidade da função máximo, mas apontando qual é o menor valor do intervalo das células.

Figura 3.26: Função *=MÍNIMO*

Desvio Padrão

O cálculo do desvio padrão, é dado pela seguinte função *=DESVPAD*, o Excel calcula o desvio padrão de tanto de amostras quanto de populações, variando de sua fórmula de *=DESVPAD.P* para populações, e *=DESVPAD.A* para amostras.

	A	B
1	comprimento(cm)	=DESVPAD.P(A2:A52)
2	10	DESVPAD.P(núm1; [núm2];
3	11,2	
4	7,9	
5	8,1	
6	9,5	
7	10,1	

Figura 3.27: Mostra a função *=DESVPAD.P* para populações

	A	B
1	comprimento(cm)	=DESVPAD.A(A2:A7)
2	10	DESVPAD.A(núm1; [núm2];
3	11,2	
4	7,9	
5	8,1	
6	9,5	
7	10,1	

Figura 3.28: Mostra a função *=DESVPAD.A* para amostras

Variância

A variância calcula a variação em relação à média de uma série de dado, dado pela seguinte fórmula *=VAR*, e assim como o desvio padrão, o Excel tem fórmula para população e amostras, *=VAR.P* para população e *=VAR.A* para a amostras.

	A	B
1	comprimento(cm)	=VAR.P(A2:A52)
2	10	VAR.P(núm1; [núm2];
3	11,2	
4	7,9	
5	8,1	
6	9,5	
7	10,1	

Figura 3.29: Mostra a função *=VAR.P* para populações

	A	B
1	comprimento(cm)	=VAR.A(A2:A7)
2	10	VAR.A(núm1; [núm2];
3	11,2	
4	7,9	
5	8,1	
6	9,5	
7	10,1	

Figura 3.30: Mostra a função *=VAR.A* para amostras

Erro Padrão

O Erro Padrão é uma medida da estimativa do intervalo de confiança da média.

(3.1): $$EP = \frac{Desvio\ Padrão}{\sqrt{N^\circ\ de\ Amostras}}$$ Fórmula do Erro Padrão

No Excel não há uma função específica. Mas é bastante simples, transcrever a fórmula no Excel da seguinte forma.

Para Calcular o Erro Padrão, deve usar a função *=RAIZ*, para elevar o número de amostras a raiz quadrada, e fazer a divisão a **/** (barra).

	A	B	C
1	10	Desvio Padrão	1,265965
2	11,2	N°de Amostras	6
3	7,9	Erro Padrão	=C1/RAIZ(C2)
4	8,1		
5	9,5		
6	10,1		

Figura 3.31: Exemplo de cálculo do Erro Padrão no Excel

Nota: C1 é o Desvio Padrão, Divisão (/), RAIZ e raiz quadrada e C2 é o número de amostras.

Coeficiente de Variação

O Coeficiente de Variação é a dispersão total da amostra expressa em porcentagem, calculada da seguinte fórmula:

(3.2): $$CV = \frac{\text{Desvio Padrão}}{\text{Média}} \times 100$$ Fórmula do Coeficiente de Variação

No Excel não há uma função específica. Mas é bastante simples, transcrever a fórmula no Excel da seguinte forma (Figura 3.32).

	A	B	C	D
1	10	Desvio Padrão	1,265965	
2	11,2	Média	9,466667	
3	7,9	Coef. de Variação	=(C1/C2)*MULT(100)	
4	8,1			MULT(núm1;
5	9,5			
6	10,1			

Figura 3.32: Exemplo de cálculo do Coeficiente de Variação no Excel.

Onde; C1 é o Desvio Padrão, e C2 é a Média.

Nota: Ao realizar a multiplicação por 100, pode-se utilizar a função =MULT ou o sinal de * (Asterisco).

Tamanho Da Amostra

No Excel para saber o número de amostras utiliza-se a função *=CONT.NÚM*, basta selecionar todos os elementos da amostra.

	A	B	C	D	E
1	10				
2	11,2	N°de Amostras	=CONT.NÚM(A1:A6)		
3	7,9		CONT.NÚM(valor1; [valor2]; ...)		
4	8,1				
5	9,5				
6	10,1				

Figura 3.33: Exemplo de cálculo do Coeficiente de Variação no Excel

GRÁFICOS

Gráficos é uma forma fácil de representar as informações e dados contidos em uma planilha. É um recurso muito funcional por proporciona uma melhor visualização das informações inseridas.

Tipos de Gráficos

O Excel tem suporte a vários modelos de gráficos, cada modelo de gráfico se subdivide, possibilitando uma grande variedade de gráficos para criar ou modificar.

Os principais tipos de gráficos são:

- **COLUNAS:**
- **LINHAS:**
- **BARRAS:**
- **PIZZA:**
- **DISPERSÃO:**
- **HISTOGRAMA**

O Excel tem a capacidade de gerar gráficos automaticamente com os dados selecionados, ⇒ *Inserir,* ⇒ *Escolher o Tipo de Modelo do Gráfico***,** e o Excel irá inserir o gráfico escolhido dos dados selecionando, mas em algumas situações o Excel não conseguirá compreender os dados, inserindo um gráfico desconfigurado.

Nota:

- Para fins educacionais os gráficos serão criados inserindo os dados manualmente.
- Como os passos iniciais para criar diferentes tipos de gráficos são semelhantes, será apresentado unicamente todos os passos para criar os gráficos de colunas.

Gráfico de Colunas

O gráfico de colunas exibe as alterações dos dados em um período de tempo ou ilustra comparações entre itens. As categorias são organizadas horizontalmente, e os valores, verticalmente.

Para criar um gráfico de colunas inserindo os dados manualmente basta conhecer os valores das médias dos fatores que serão comparados, e para acrescentar a barra de erro, basta calcular o desvio padrão ou erro padrão.

PASSO 1: Para inserir um gráfico de colunas com os dados da **Tabela 3.1**, precisa calcular a média de cada tipo de sedimento. Primeiramente os dados serão transferidos para o Excel.

AREIA	AREIA LAMOSA	CASCALHO	CASCALHO ARENOSO
34	26	37	23
26	37	45	28
33	42	39	30
36	34		37
	36		32

Tabela 3.1: Biomassa de Camarão em função do tipo de sedimento

	Biomassa de Camarão por arrasto em função do tipo de sedimento			
Tipos de Sedimento	AREIA	AREIA LAMOSA	CASCALHO	CASCALHO ARENOSO
	34	26	37	23
	26	37	45	28
	33	42	39	30
	36	34		37
		36		32
				29

Figura 3.34: Dados da Tabela 3.1, organizados no Excel

PASSO 2: Após transferir os dados para Excel, será calculado média e o desvio padrão, com as funções *=MÉDIA* e *=DESVPAD.A* do Excel.

Utiliza as funções de *=MÉDIA* e *=DESVPAD.A*, basta selecionar o intervalo de valores da primeira coluna *(AREIA),* **Figura 3.35.**

Tipos de Sedimento	AREIA	AREIA LAMOSA
	34	26
	26	37
Área do Gráfico	33	42
	36	34
		36
Média	=MÉDIA(C5:C8)	

Figura 3.35: Cálculo da média da Tabela 3.1

PASSO 3: Para calcular a média e o desvio padrão das demais colunas com *Autopreenchimento*, basta colocar o cursor do mouse no canto inferior direito da célula com resultado da primeira coluna, quando o cursor mouse mudar para uma cruz, basta arrastar

até a última coluna, **Figura 3.36.**

AREIA	AREIA LAMOSA	CASCALHO	CASCALHO ARENOSO
34	26	37	23
26	37	45	28
33	42	39	30
36	34		37
	36		32
			29
32,25			

Figura 3.36: Cálculo com Autopreenchimento

Após isso o Excel irá calcular automaticamente as linhas de cada coluna (*tipos de sentimentos*) da planilha, **Figura 3.37.**

Tipos de Sedimento	AREIA	AREIA LAMOSA	CASCALHO	CASCALHO ARENOSO
	34	26	37	23
	26	37	45	28
	33	42	39	30
	36	34		37
		36		32
				29
Média	32,25	35	40,3333	29,83333333

Figura 3.37: Cálculo da média finalizado

PASSO 4: Após calcular a média e o desvio de padrão, selecionar uma célula em branco, ⇒ guia *Inserir,* ⇒ *Inserir Gráfico de Colunas ou Barras*, depois na primeira opção ⇒ (*Colunas Agrupadas*), **Figura 3.38.**

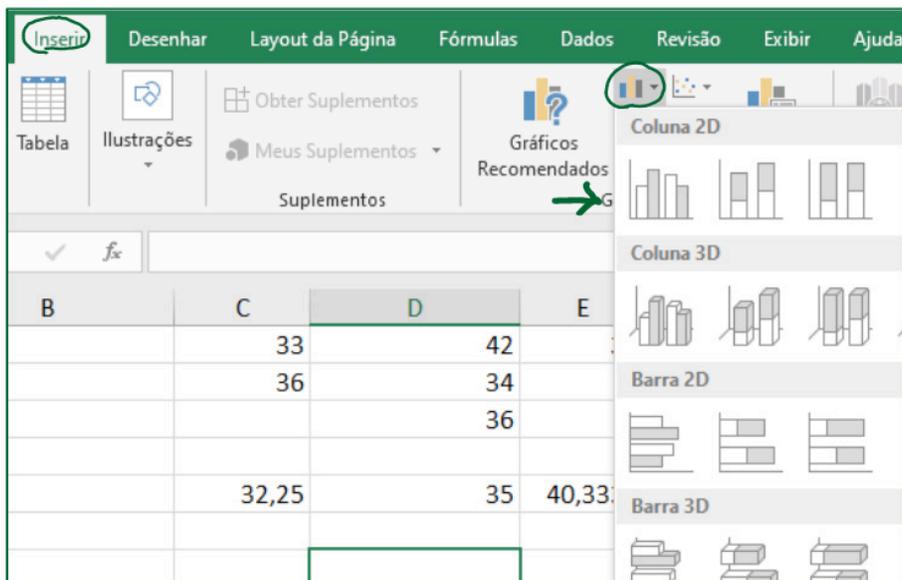



Figura 3.38: Inserindo o Gráfico de Colunas

PASSO 5: O Excel vai criar um gráfico em brando, para adicionarmos os valores, após isso na nova guia que surgiu, ⇒ *Design,* ⇒ *Selecionar Dados,* **Figura 3.39.**

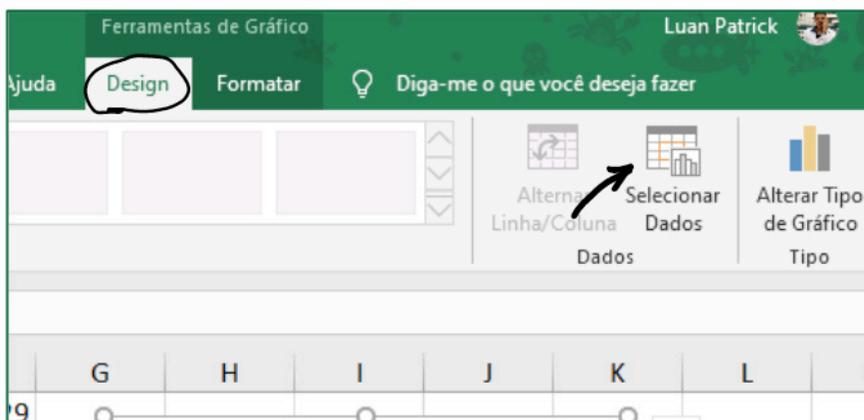



Figura 3.39: Selecionando os Dados para o gráfico

PASSO 6: Uma nova janela vai se abrir (*Selecionar Fonte de Dados*), ⇒ *Adicionar,* ⇒ *Valores da série*, selecione as células com os resultados da média, clique em *OK*. **Figura 3.40.**

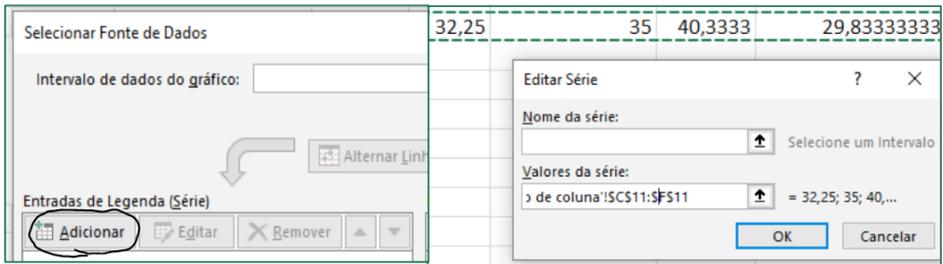



Figura 3.40: Dados selecionados

PASSO 7: Novamente na janela *Selecionar Fonte de Dados*, em *Rótulos do eixo Horizontal*, clicar em ⇒ *Editar*, para mostrar ao Excel o que cada coluna irá representar, **Figura 3.41.** O Excel por sua vez, vai inserir rótulos nas colunas do gráfico, conforme os títulos da séries na tabela, no caso da **Tabela 3.1**, é sedimentos, o título das colunas será os diferentes tipos de sedimentos.

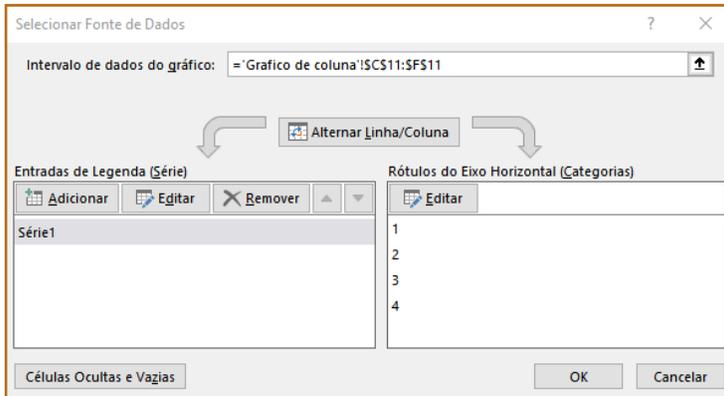



Figura 3.41: Inserindo Rótulos nas Colunas

PASSO 8: Na janela *Rótulos do Eixo,⇒ selecionar os intervalos de células que contém os títulos das séries*, após isso ⇒ *clicar no botão* **Figura 3.42.**

dimen	AREIA	AREIA LAMO:	CASCALH	CASCALHO ARENOSO
	34	26	37	23
	26	37	45	28
	33	42	39	30
	36	34		37

Rótulos do Eixo ? ×

='Grafico de coluna'!$C$4:$G$4

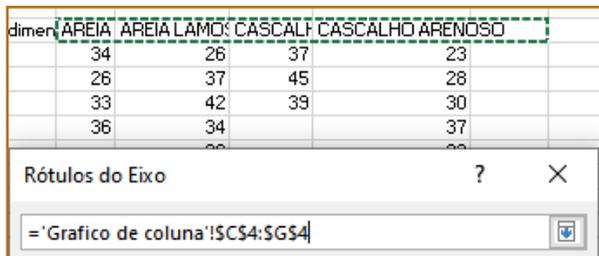



Figura 3.42: Selecionando os títulos dos rótulos dos eixos

PASSO 9: Na janela *Selecionar Fonte de Dados*, o Excel terá reconhecido os títulos das séries, para finalizar clicar em ⇒ *OK*, **Figura 3.43.**

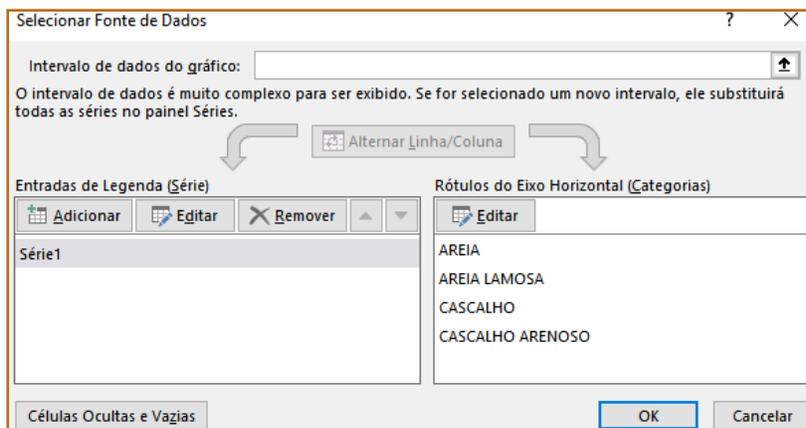



Figura 3.43: Finalizando a inserção do Gráfico

O Excel irá adicionar as colunas e o títulos das colunas no gráfico, **Figura 3.44.**

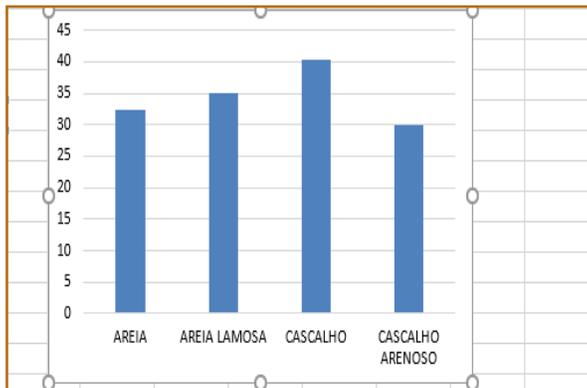



Figura 3.44: Gráfico com os dados da **Tabela 3.1**

PASSO 10: Para incluir no gráfico de colunas a Barra de Erro primeiro deve-se selecionar o gráfico recém criando, ⇒ *Elementos de Gráfico* (Botão verde), ⇒ *Barras de Erros,* ⇒ *Mais Opções*, **Figura 3.45.**

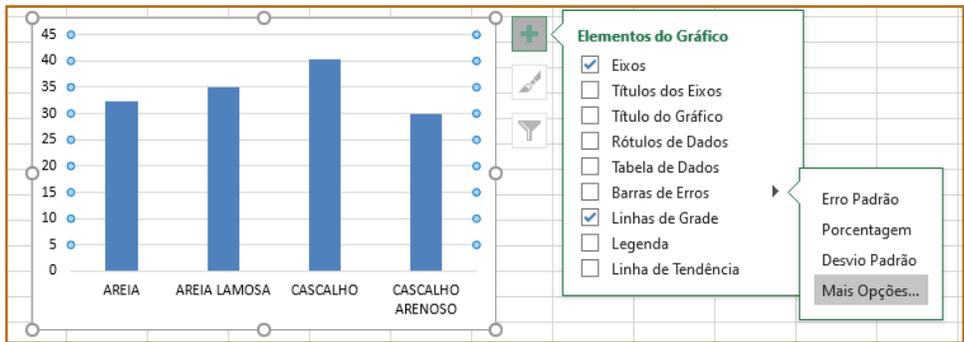



Figura 3.45: Incluindo Barra de Erro

PASSO 11: Na nova janela que se abriu ⇒ *Na última opção* da *área Erro,* ⇒ *Personalizando***,** ⇒ clicar em *Especificar Valor*. Na janela *Barras de Erros Personalizadas*, ⇒*Valor do Erro Positivo* e *Valor de Erro Negativo* ⇒ *Selecionar os Valores do Desvio Padrão*, **Figura 3.46.**

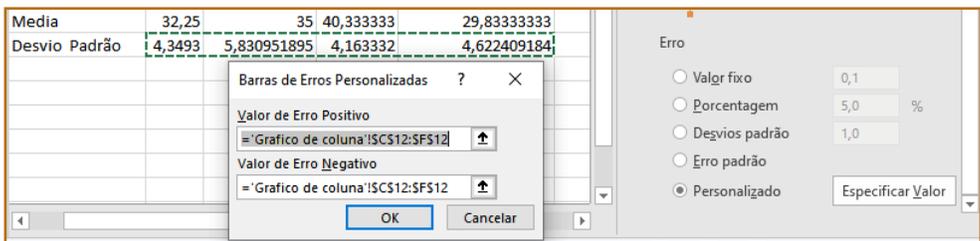



Figura 3.46: Seleção dos Valores da Barra de Erro

PASSO 12: No gráfico não está especificado o título do eixo dos valores (Y), para especificar, ⇒ *Elementos de Gráfico* (Botão verde), ⇒ *Títulos dos Eixos,* ⇒ Renomeie os títulos dos eixos. Como na **Tabela 3.1** compara a biomassa, o título dos valores é Biomassa e a unidade de medida da biomassa, **Figura 3.47.**

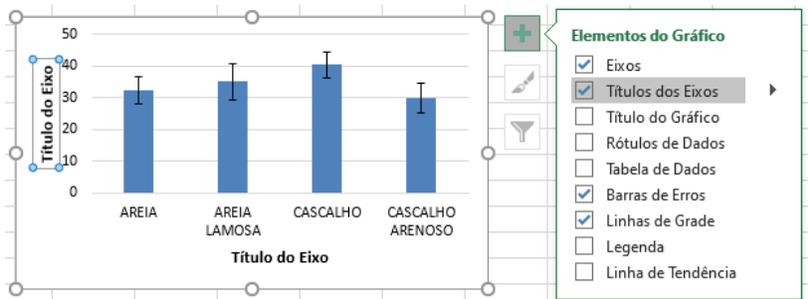



Figura 3.47: Inserindo Títulos nos Eixo

O gráfico de coluna da **Figura 3.48** está apto para a publicação em revistas e trabalhos científicos. Há outras configurações para modificar a aparência, mas não será abordado.

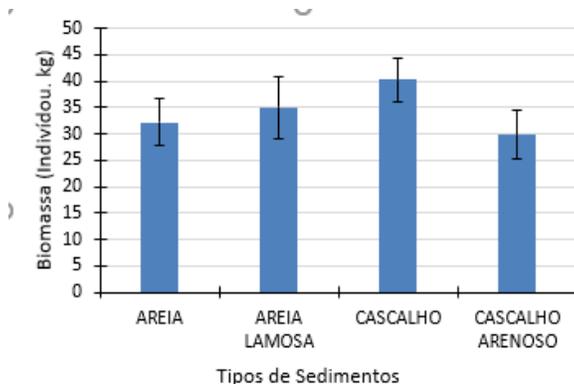



Figura 3.48: Gráfico de Colunas da Tabela 3.1 finalizando.

Gráfico de Barras

Para inserir um gráfico de barras com os dados da **Tabela 3.2** as linhas é o eixo das séries os serão comparados (Espécies de Peixes pescados no município de Santarém), e as Colunas é o eixo dos valores, são as frequências de ocorrência dos valores (Quilos de Peixes capturados).

Espécies	JAN.	FEV.	MAR.	ABR.	MAI.	JUN.
Aracú	9	103	4212	7066	4761	3273
Dourada	1248	1828	1353	2055	2051	3235
Jaraqui	4363	5656	2162	7427	15905	10194
Mapará	0	0	4448	8476	2598	5349
Pacú	255	0	1246	5889	7987	3319
Pescada	4635	3016	5216	4245	1024	2642
Surubim	4584	5213	3251	3352	2501	1811
Tambaqui	216	235	3038	5806	6328	4475
Tucunaré	5295	3852	1763	643	766	2121

Tabela 3.2: Biomassa de Camarão em função do tipo de sedimento

PASSO 1: Inserir dados da **Tabela 3.2 n**o Excel.

	A	B	C	D	E	F	G	H	I
2		Produção total de Pescado no munícipio de Santarém em quilos (kg)							
4		Espécies	JAN.	FEV.	MAR.	ABR.	MAI.	JUN.	
5		Aracú	9	103	4212	7066	4761	3273	
6		Dourada	1248	1828	1353	2055	2051	3235	
7		Jaraqui	4363	5656	2162	7427	15905	10194	
8		Mapará	0	0	4448	8476	2598	5349	
9		Pacú	255	0	1246	5889	7987	3319	
10		Pescada	4635	3016	5216	4245	1024	2642	
11		Surubim	4584	5213	3251	3352	2501	1811	
12		Tambaqui	216	235	3038	5806	6328	4475	
13		Tucunaré	5295	3852	1763	643	766	2121	

Figura 3.49: Inserindo dados da Tabela 3.2 no Excel

PASSO 2: Adicionar Nova coluna (após a coluna *JUN.*), ⇒ com a função *=SOMA* para somar todos os valores a primeira linha *(Aracú*), ⇒ *Autopreenchimento* calcular as demais linhas (*séries*) **Figura 3.50.**

	A	B	C	D	E	F	G	H	I	J
2		Produção total de Pescado no munícipio de Santarém em quilos (kg)								
4		Espécies	JAN.	FEV.	MAR.	ABR.	MAI.	JUN.	Total	
5		Aracú	9	103	4212	7066	4761	3273	=SOMA(C5:H5)	
6		Dourada	1248	1828	1353	2055	2051	3235	11770	
7		Jaraqui	4363	5656	2162	7427	15905	10194	45707	
8		Mapará	0	0	4448	8476	2598	5349	20871	
9		Pacú	255	0	1246	5889	7987	3319	18696	
10		Pescada	4635	3016	5216	4245	1024	2642	20778	
11		Surubim	4584	5213	3251	3352	2501	1811	20712	
12		Tambaqui	216	235	3038	5806	6328	4475	20098	
13		Tucunaré	5295	3852	1763	643	766	2121	14440	
14										

SOMA(núm1; [nú

Figura 3.50: Inserindo Nova Coluna (*Total*) o somatório dos fatores

PASSO 3: Selecionar todos os elementos da tabela, ⇒ guia *Página Inicial,* ⇒ *Classificar e Filtrar,* ⇒ *Classificação Personalizada,* ⇒ *Classificar por "Total",* ⇒ *Ordem Do Menos para o Maior*. **Figura 3.51.**

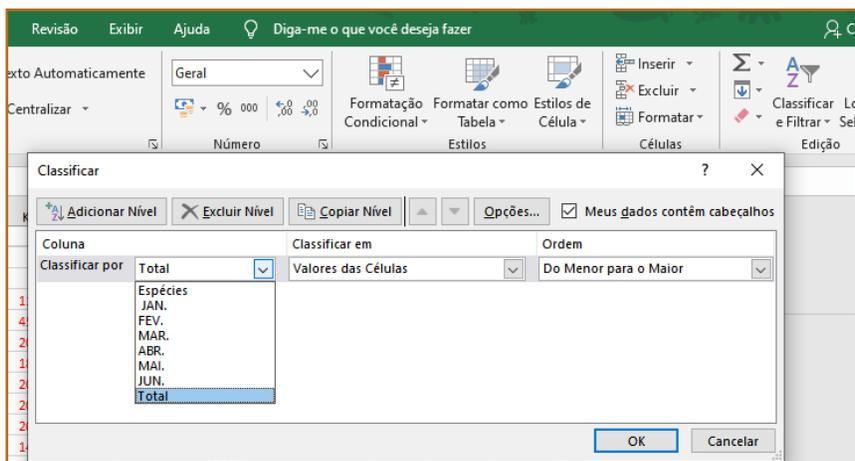



Figura 3.51: Organizando em Ordem de Valores

PASSO 4: Selecionar *Célula em branco* ⇒ *Inserir,* ⇒ *Inserir Gráfico de Colunas ou de Barras,* ⇒ *Barras Agrupas.*

PASSO 5: ⇒ *Selecionar Dados,* ⇒ *Adicionar,* ⇒ *Valores da série*, ⇒ Selecionar os dados da coluna "*Total*" ⇒ **Ok.**

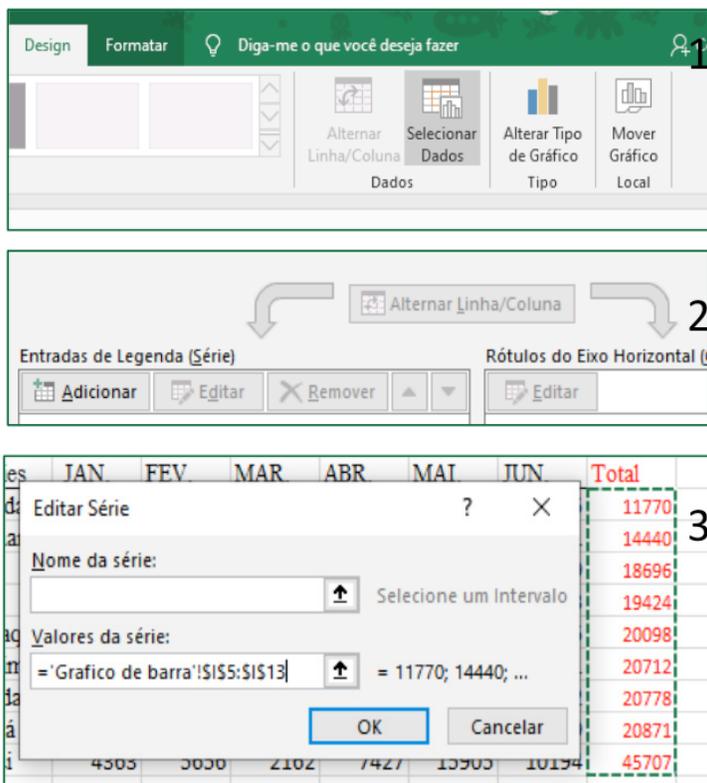



Figura 3.52: Inserindo Valores da Série

PASSO 6: Em *Rótulos do Eixo Horizontal* ⇒ *Editar,* ⇒ *Intervalo do rótulo do eixo,* ⇒Selecionar os dados da coluna "*Espécies*", ⇒ *OK* ⇒ *OK.*

Figura 3.53: Inserindo Rótulos do Eixo

PASSO 7: Para adicionar o Títulos dos Eixos repetir o **PASSO 12** do gráfico de colunas, com a diferença que o gráfico de Barra investe eixo da série e dos valores.

O resultado do gráfico de barras da **Tabela 3.2** na **Figura 3.54.**

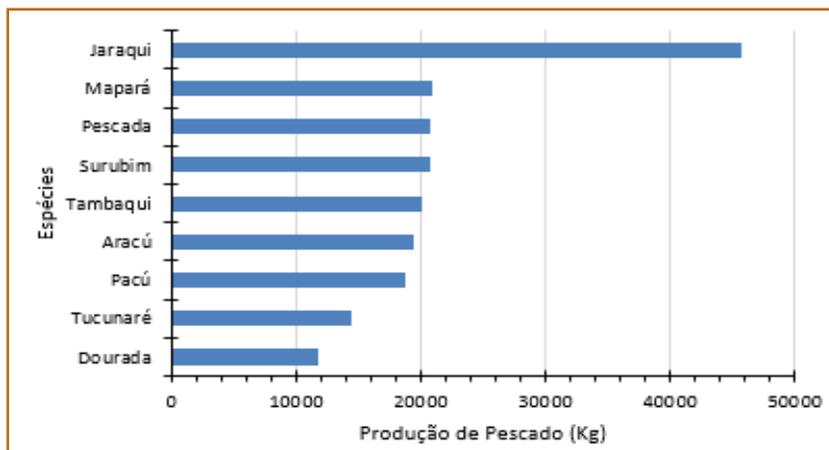



Figura 3.54: Gráfico de Barras da Tabela 3.2 finalizado

Histograma

É um gráfico de colunas que demonstra a distribuição em frequências das amostras, onde o eixo horizontal são os intervalos dos valores, e o eixo vertical é a frequência do número de vezes que os intervalos foram observados.

PASSO 1: Inserir os dados no Excel.

Será utilizando como exemplo dados de 60 amostras de peso(g) de camarão da Amazônia, *Macrobrachium amazonicum,* organizados em uma única coluna.

PASSO 2: Classificar os valores das amostras em ordem *Do Menor para o Maior (Selecione os dados ⇒ Classificar e Filtrar ⇒ Classificação Personalizada ⇒ Do Menos para o Maior).*

PASSO 3: Calcular o *Número de Amostras*, utiliza-se a função *=CONT.NÚM.* No exemplo são 60 amostras de peso de camarão.

PASSO 4: Calcular *o Número de Classes (M),* **Figura 3.55**, O número de Classes define em quantos intervalos os dados da amostra serão divididos. É definida pela fórmula:

$$\mathrm{M} = 0{,}9 \times \sqrt{n} \quad (3.4)$$

Onde **0,9** é uma constante, n número de amostras.

E1 | =0,9*RAIZ(E1)

	A	B	C	D	E	
1	Peso (g)		N	Número de amostra	60	
2	12,4		M	Número de Classes	=0,9*RAIZ(E1)	7 ←
3	13,1					

Figura 3.55: Cálculo do Número de Classes no Excel

O resultado do cálculo do número de classes, deve arredondado para um número inteiro acima ou abaixo do valor encontrado. Como resultado da **Figura 3.55** foi 6,97 assume-se o valor *7,* preencher a célula ao lado com o valor arrendado 7.

PASSO 5: Calcular a *Amplitude* das amostras, primeiro encontre o valor máximo e o valor mínimo com a função *=MÁXIMO* e *=MÍNIMO*, ⇒ *Fórmula da Amplitude (R).*

(3.5): $$R = MÁXIMO - MÍNIMO$$

PASSO 6: Calcular o *Intervalo de Classe (I)* com a fórmula:

(3.6): $$I = \frac{R}{M}$$

Onde R é *a Amplitude*, M é o *Número de Classes*.

	A	B	C	D	E	F
1	Peso (g)		N	Número de amostra	60	60
2	12,4		M	Número de Classes	6,971	7
3	13,1		R	Amplitude	14	14
4	14,5		I	Intervalo de Classe	=E3/E2	2 ←

Figura 3.56: Cálculo do Intervalo de Classe no Excel

No exemplo o resultado do cálculo do intervalo de classe apresentado na **Figura 3.56** foi de 2,01, arredonde para o valor 2, preencha a célula ao lado com o valor arredado 2.

PASSO 6: Construir uma tabela semelhante a **Figura 3.57** com os *limites de Classe* e as *frequências Absolutas* e *Relativas*.

Classes	Limite Inferior	Limite Superior	Ponto Médio	Fa	Fr(%)
1					
2					
3					
4					
5					
6					
7					

Figura 3.57: Tabela do Limites de Classes

Na tabela da **Figura 3.57** as *Classes* são os números de classes (No cálculo do Número de Classes foi valor 7, então o serão 7 classes), *Limite Inferior* determina o valor inicial do intervalo (intervalo fechado), *Limite Superior* de delimita o valor extremo do final do intervalo (Intervalo Aberto). *Ponto Médio* é a média dos valores inferior e superior, ***Fa*** é a *Frequência Absoluta* dos intervalos, ***Fr****(%)* é a *Frequência da Relativa* dos intervalos.

Para o preenchimento da tabela **Figura 3.57** é importante utilizar o modo autopreenchimento do Excel utilizando o *sinal de igual (=)* e ⇒ *clicando na Célula o valor arrendado* para facilitar nessa tarefa.

PASSO 7: Calcular o *Limite Inferior* e *Superior* das classes das amostras na tabela apresentada na **Figura 3.57**. ⇒ Na *Classe 1* coluna *Limite Inferior* preencher com o *Valor Mínimo* das amostras (12,4),⇒ Na próxima coluna *Limite Superior* realizar o seguinte cálculo:

(3.7): Limite Superior = Limite Inferior + Intervalo de Classe

C	D	E	F	H	I	J
N	Número de amostra	60	60			
M	Número de Classes	6,971	7	Classes	Limite Inferior	Limite Superior
R	Amplitude	14	14	1	12,4	=I3+F4
I	Intervalo de Classe	2,01	2	2		

Figura 3.58: Cálculo do *1º Limite Superior*

No exemplo o *Limite Inferior* (12,4) **+** (mais) o *Intervalo de Classe* (2) resultando no *Limite Superior* de 14,4.

PASSO 8: Na coluna *Limite Inferior* da *Classe 2*, preencher com o resultado do *Limite Superior* da *Classe 1* (14,4).

D	E	F	H	I	J
Número de amostra	60	60			
Número de Classes	6,971	7	Classes	Limite Inferior	Limite Superior
Amplitude	14	14	1	12,4	14,4
Intervalo de Classe	2,01	2	2	=J3	

Figura 3.59: Cálculo do *2º Limite Inferior*

PASSO 9: Na coluna *Limite Superior* da *Classe 2,* preenche com o cálculo do *Limite Inferior* (14,4) **+**(mais) o *Intervalo de Classe* (2). colocar a *barra de digitação* entre a letra e o número da célula e pressionar a *tecla F4*, para ancorar o valor 2.

D	E	F	H	I	J	K
Número de amostra	60	60				
Número de Classes	6,971	7	Classes	Limite Inferior	Limite Superior	Pont
Amplitude	14	14	1	12,4	14,4	
Intervalo de Classe	2,01	2	2	14,4	=I4+F4	

Figura 3.60: Cálculo do 2º *Limite Superior*

PASSO 10: Para calcular os *Limites Inferior e Superior* das outras *Classes,* ⇒ Primeiro precisa-se fixar o valor do *Intervalo de Classe* colocando a *barra de digitação* entre a *letra* e o *número* do endereço da célula, ⇒ e pressionar a *tecla F4,* para ancorar o valor 2 do *Intervalo de Classe*.

D	E	F	H	I	J	
Número de amostra	60	60				
Número de Classes	6,971	7	Classes	Limite Inferior	Limite Superior	P
Amplitude	14	14	1	12,4	14,4	
Intervalo de Classe	2,01	2	2	14,4	=I4+F4	
			3			

D	E	F	H	I	J	
Número de amostra	60	60				
Número de Classes	6,971	7	Classes	Limite Inferior	Limite Superior	Pe
Amplitude	14	14	1	12,4	14,4	
Intervalo de Classe	2,01	2	2	14,4	=I4+$F$4	

Figura 3.61: Fixando o valor do *Intervalo de Classe*

Após pressionar a *tecla F4* o endereço da célula irá r mudar aparecendo dois $ (cifrão) na frente da *letra* e do *número*, exemplo: A1 ⇒ $A$1

PASSO 11: Selecionar o *Limite Inferior* e *Superior* da *Classe 2*, utilizando o *autopreenchimento* arrastar até a última *Classe* (*Classe* 7) para calcular os *Limites Inferiores e Superiores* das demais *Classes*.

7	Classes	Limite Inferior	Limite Superior	Ponto	Fa	Fr(%)
14	1	12,4	14,4			
2	2	14,4	16,4			
	3					
	4					
	5					
	6					
	7					

Figura 3.62: Calculando os *Limite das Classes*

O autopreenchimento irá calcular automaticamente os limites das próximas *Classe*. No exemplo o *Limite Inferior* foi 24,4 e o *Limite Superior* foi 26,4 da *Classe 7.*

PASSO 12: Calcular os Pontos Médios da *Classe 1* (O *Ponte Médio* é a média do *Limite Inferior* e *Limite Superior* de cada Classes). Utilizar o autopreenchimento para

calcular o *Ponto Médio* das demais *Classes*.

Classes	Limite Inferior	Limite Superior	Ponto Médio	Fa
1	12,4	14,4	=(I3+J3)/2	

Figura 3.63: Cálculo do Ponto Médio

No cálculo do *Ponto Médio* pode-se utilizar a função *=MÉDIA* ou somar os limites e dividir por 2.

PASSO 13: Na coluna *Fa* (*Frequência Absoluta*) preencher com número de observações das amostras nos intervalos de cada *Classe*. Importante lembrar que o intervalo do *Limite Inferior* é fechado e do *Limite Superior* é Aberto, isso significa que o intervalo começa no *Limite Inferior* e termina próximo ao *Limite Superior*, pois o *limite Superior* de uma *Classe* será o próximo *Limite Inferior* da *Classe* seguinte.

No exemplo *na Classe 1* o *Limite Inferior* é 12,4 e o *Limite Superior* é 14,4 a *Frequência Absoluta* irá ser o número de observações das amostras aparecem no intervalo de 12,4 até próximo de 14,4. Na *Classe 2* o *Limite Inferior* é 14,4 e o *Limite Superior* é 16,4.

Classes	Limite Inferior	Limite Superior	Ponto Médio	Fa	Fr(%)
1	12,4	14,4	13,4	2	
2	14,4	16,4	15,4	4	
3	16,4	18,4	17,4	9	
4	18,4	20,4	19,4	28	
5	20,4	22,4	21,4	6	
6	22,4	24,4	23,4	8	
7	24,4	26,4	25,4	3	
			Total	60	

Figura 3.64: Cálculo *Frequência Absoluta*

A frequência absoluta de cada Classe foi calculada com manualmente contando os números de observações das amostras nos intervalos. Exemplo, se amostra for de 17,4 segundo a **Figura 3.64** irá se encaixar no intervalo de 16,4 a 18,4 da *Classe 3*.

O somatório da *Frequência Absoluta* será o número de total amostras. No exemplo são 60 amostras, então o somatório da frequência absoluta será 60.

PASSO 14: Calcular a ***Fr*** (*Frequência Relativa*) de cada *Classe* dividindo *Frequência Absoluta* das *Classes* pelo total das amostras.

Classes	Limite Inferior	Limite Superior	Ponto Médio	Fa	Fr(%)
1	12,4	14,4	13,4	2	=L3/L10
2	14,4	16,4	15,4	4	
3	16,4	18,4	17,4	9	
4	18,4	20,4	19,4	28	
5	20,4	22,4	21,4	6	
6	22,4	24,4	23,4	8	
7	24,4	26,4	25,4	3	
			Total	60	

Figura 3.65:Cálculo *Frequência Relativa*

Para calcular a Frequência Relativa das próximas Classes deve-se fixar o valor total da *Frequência Absoluta* utilizando a *tecla F4* e com o *Autopreenchimento* arrastar até a última Classe.

Classes	Limite Inferior	Limite Superior	Ponto Médio	Fa	Fr(%)
1	12,4	14,4	13,4	2	0,033
2	14,4	16,4	15,4	4	0,067
3	16,4	18,4	17,4	9	0,15
4	18,4	20,4	19,4	28	0,467
5	20,4	22,4	21,4	6	0,1
6	22,4	24,4	23,4	8	0,133
7	24,4	26,4	25,4	3	0,05
			Total	60	

Figura 3.66:Cálculo Frequência Relativa

PASSO 15: Selecionar os valores da coluna ***Fr*** (*Frequência Absoluta*) ⇒ Na guia *Página Inicial*, ⇒ faixa *Número*, ⇒ marcar a opção *Porcentagem* ***(%)*** ou pressionar as teclas *Ctrl+Shif+%,* para deixar os valores da coluna em porcentagem.

F	H	I	J	K	L	M
60						
7	Classes	Limite Inferior	Limite Superior	Ponto Médio	Fa	Fr(%)
14	1	12,4	14,4	13,4	2	3%
2	2	14,4	16,4	15,4	4	7%
	3	16,4	18,4	17,4	9	15%
	4	18,4	20,4	19,4	28	47%
	5	20,4	22,4	21,4	6	10%
	6	22,4	24,4	23,4	8	13%

Figura 3.67:Cálculo Frequência Relativa

PASSO 16: Inserir o gráfico do histograma, ⇒ *clicar* em cé*l*ula em *branco*, ⇒ guia *Inserir*, ⇒ *inserir de Gráfico Coluna Agrupada*, ⇒*Guia Design*, ⇒ *Selecionar Dados,* ⇒ *Adicionar,* ⇒ em *Valores da série*,⇒ *seleciona*r os valores da *Frequência Absoluta* (*Fa*), ⇒ *Ok.*

Classes	Limite Inferior	Limite Superior	Ponto Médio	Fa	Fr(%)
1	12,4	14,4	13,4	2	3%
2	14,4	16,4	15,4	4	7%
3	16,4	18,4	17,4	9	15%
4	18,4	20,4	19,4	28	47%
5	20,4	22,4	21,4	6	10%
6	22,4	24,4	23,4	8	13%
7	24,4	26,4	25,4	3	5%
			Total	60	

Figura 3.68: Inserindo Gráfico de Coluna

PASSO 17: Em *Rótulos do Eixo Horizontal, Editar*, ⇒ em *Intervalo de rótulo* selecionar os valores do *Ponto Médio*, ⇒ *Ok.*

erior	Limite Superior	Ponto Médio	Fa	Fr(%)
12,4	14,4	13,4	2	3%
14,4	16,4	15,4	4	7%
16,4	18,4	17,4	9	15%
18,4	20,4	19,4	28	47%
20,4	22,4	21,4	6	10%
22,4	24,4	23,4	8	13%
24,4	26,4	25,4	3	5%
		Total	60	

Rótulos do Eixo

Intervalo do rótulo do eixo:

=Histograma!$K$3:$K$9 = 13,4; 15,4; 17...

OK Cancelar

Figura 3.69: Especificando os Rótulos do Eixo

PASSO 18: Adicionar *Segunda Série*, ⇒ *Adicionar*, ⇒ em *Valores* da *série*, selecionar os valores da *Frequência Relativa* (*Fr*), ⇒ *Ok*, ⇒ *Ok*.

Figura 3.70: Gráfico

PASSO 19: Pressionar com o *botão direito* no gráfico, ⇒ *Alterar Tipo Gráfico*, ⇒ Em *Gráfico Combinação*, ⇒ Selecionar *Coluna Clusterizada – Linha no Eixo Secundário*

Na Série2 deixe em *Tipo de Gráfico* deixar em *Linha* e marcar a opção *Eixo Secundário*, como apresentado no **Figura 3.71.**

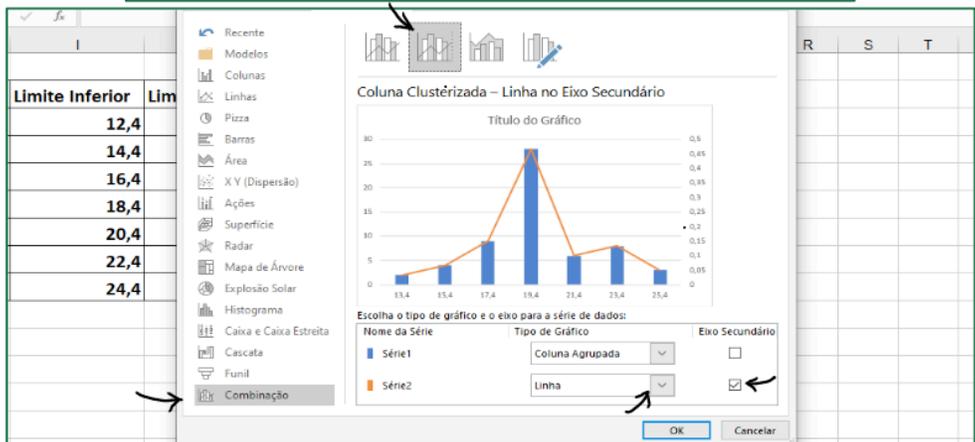



Figura 3.71: Eixo Secundário *Frequência Relativa*

PASSO 20: ⇒ *Elemento do Gráfico*, ⇒ Adicionar *Títulos dos Eixos*, ⇒ *Renomeie os Eixos semelhante a* **Figura 3.72.**

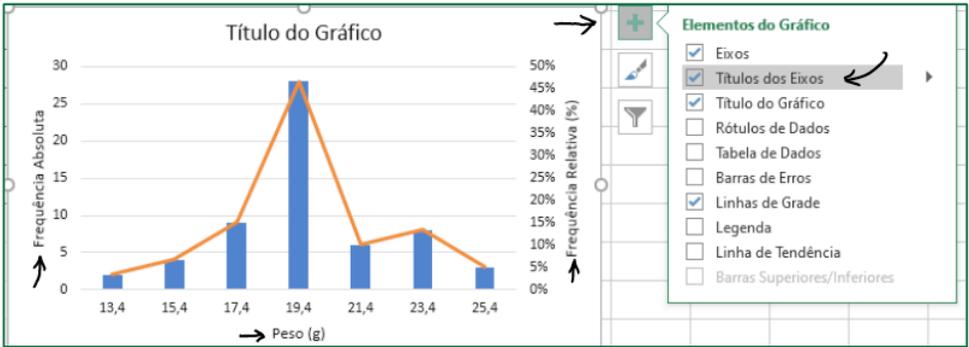



Figura 3.72: Adicionando o Títulos dos Eixos

PASSO 21: ⇒ *Dois Cliques* na coluna do gráfico, ⇒ em *Opções de Série*, Marcar *Eixo Principal Largura do Espaçamento Diminuir* a porcentagem para 2%, para aumentar a

largura das colunas e diminuir o espaçamento.

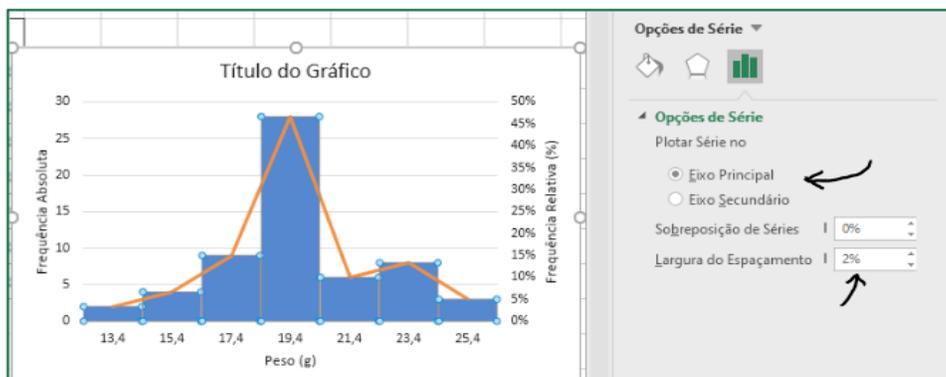



Figura 3.73: Aumentando Largura das Colunas

A **Figura 3.73** apresenta o Histograma finalizado. Preparado para publicações em revistas e apresentações científicas.

Figura 3.74: Histograma Finalizado

TABELAS

As tabelas de dados para análise sempre devem ser construídas com os dados ou variáveis nas colunas e as unidades amostrais, sujeitos ou indivíduos nas linhas. Basicamente, há dois formatos de tabelas que podem ser construídos: o formato largo e o formato longo.

No formato largo, cada nível do fator ocupa uma coluna diferente, seria como se cada nível fosse uma variável diferente, mas deve-se ter em mente que eles são apenas níveis diferentes do mesmo fator ou da mesma variável. No formato longo, o fator ocupa

apenas uma coluna e os níveis são especificados nas linhas.

Réplicas	GrupoA	GrupoB	GrupoC
1	2,2	5,2	6,4
2	4,6	4,3	9,9
3	3,1	6,3	8,6
4	5,1	5,8	6,3
5	4,7	5,1	7,6

Tabela 3.3: Tabela no formato largo

ID	Réplicas	Grupo	Variável X
1	1	A	2,2
2	2	A	4,6
3	3	A	3,1
4	4	A	5,1
5	5	A	4,7
6	1	B	5,2
7	2	B	4,3
8	3	B	6,3
9	4	B	5,8
10	5	B	5,1
11	1	C	6,4
12	2	C	9,9
13	3	C	8,6
14	4	C	6,3
15	5	C	7,6

Tabela 3.4: Tabela no formato longo

O formato largo é mais intuitivo para o iniciante em análise de dados, já que é fácil de localizar e analisar as diferenças. Este é o formato que o Excel utiliza para produzir gráficos e possibilita calcular medidas resumo, como média ou soma, e colocar na última linha da tabela. Entretanto, é mais difícil trabalhar com este formato largo quando se tem

muitas variáveis, já que a inclusão de novos dados é mais trabalhosa e apresenta maior risco de erros, **Tabela 3.5.**

	Variável X			Variável Y		
Réplicas	**GrupoA**	**GrupoB**	**GrupoC**	**GrupoA**	**GrupoB**	**GrupoC**
1	2,2	5,2	6,4	0,7	0,7	0,7
2	4,6	4,3	9,9	0,5	0,2	0,8
3	3,1	6,3	8,6	0,1	0,6	0,8
4	5,1	5,8	6,3	0,4	0,6	0,2
5	4,7	5,1	7,6	0,7	0,5	0,4
Média	**3,94**	**5,34**	**7,76**	**0,5**	**0,5**	**0,6**

Tabela 3.5: Tabela no formato largo com duas variáveis, cada uma com três grupos.

Dados fictícios

A tabela de formato longo é preferível quando se trabalha com grande quantidade de dados, pela facilidade de incluir novas variáveis e grupos.

ID	Réplicas	Grupo	Variável X	Variável Y
1	1	A	2,2	0,7
2	2	A	4,6	0,5
3	3	A	3,1	0,1
4	4	A	5,1	0,4
5	5	A	4,7	0,7
6	1	B	5,2	0,7
7	2	B	4,3	0,2
8	3	B	6,3	0,6
9	4	B	5,8	0,6
10	5	B	5,1	0,5
11	1	C	6,4	0,7
12	2	C	9,9	0,8
13	3	C	8,6	0,8
14	4	C	6,3	0,2
15	5	C	7,6	0,4

Tabela 3.6: Tabela no formato longo com duas variáveis, cada uma com três grupos**.**

Dados fictícios

Ao construir uma tabela seja no formato largo ou longo em nenhuma circunstância colocar as unidades de medida com os valores da célula. Isto é, nunca escrever g para gramas ou cm para centímetros, dias etc. As unidades devem constar apenas no título da coluna. Exemplo: Peso_(g), altura_(cm), tempo_(dias).

Se colocar a unidade dentro da célula, o Excel não interpretará o valor como número e não será possível realizar contas, além disso, quando colar estes dados em outro programa, não será possível realizar a análise.

ATALHOS DO EXCEL

Atalhos para navegar pela planilha

Mover a tela para baixo	Page Down
Mover a tela para cima	Page Up
Mover a seleção da célula para a direita	Tab
Mover a seleção da célula para a esquerda	Shift+Tab

Atalhos para Seleção de células

Selecionar a linha inteira	Shift+Barra de espaço
Selecionar a coluna inteira	Ctrl+Barra de espaço
Selecionar todas as células da coluna com valores	Ctrl+Shift+Seta Baixo
Selecionar todas as células da linha com valores	Ctrl+Shift+Seta Direita
Expandir a seleção das células para cima	Shift+Page Up
Mover a seleção da célula para baixo	Enter
Mover a seleção da célula para cima	Shift+Enter
Mover a seleção da célula para esquerda	Shift + Tab
Cancelar a seleção	Esc

Atalhos para Inserir e Editar Dados Desfazer / Refazer

Desfazer a última ação	Ctrl+z
Refazer a última ação	Ctrl+y
Copiar Células ou conteúdo	Ctrl+C
Recortar Células ou conteúdo	Ctrl+X
Colar Células ou conteúdo	Ctrl+V

Edição Dentro Das Células

Fixar a célula	F4 entre a Letra e o Número do endereço da célula
Selecionar a Função	Tab

CAPÍTULO 4

ANÁLISE DE DADOS NO PAST

APRESENTAÇÃO DO PAST

O **PAST** é um software livre para análise de dados científicos, com funções para manipulação de dados, plotagem, estatística univariada e multivariada, análise ecológica, séries temporais e análise espacial, morfometria e estratigrafia disponibilizado pelo Øyvind Hammer.

Execução do Past

O PAST é autoexecutável e não necessita ser instalado no computador, após o realizar o download e descompactar o arquivo, basta clicar duas vezes no símbolo ( ). Por ser um arquivo com a extensão ".exe", o software de antivírus ou o sistema operacional do usuário podem perguntar se ele é confiável e pedir permissão para a execução do software. Somente após fornecer esta autorização poderá usar o software.

Interface do Past

O Past possui uma interface de usuário em formato de planilha similar ao Excel. Os dados são inseridos em uma matriz de células, organizada em linhas e colunas.

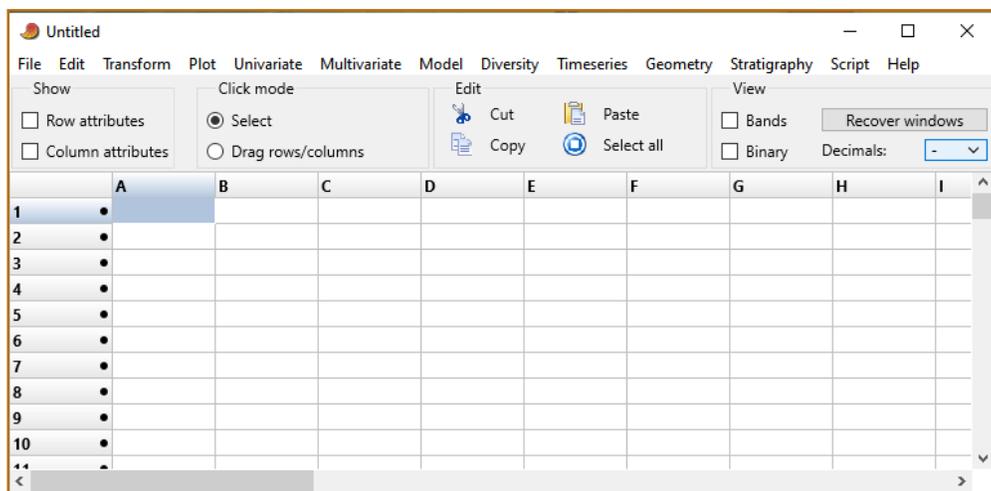



Figura 4.1: Tela inicial do programa PAST versão 4.06b (2021).

Área de Trabalho do Past

O ambiente de trabalho do Past é formado pela área da matriz de células, onde os dados são inseridos e pela região de edição e análise do Past, onde as funções do Past são organizadas em Abas. Ao se clicar em uma Abas abre as funções e os comandos associados ao nome da aba, semelhante ao Excel. **Figura 4.2.**

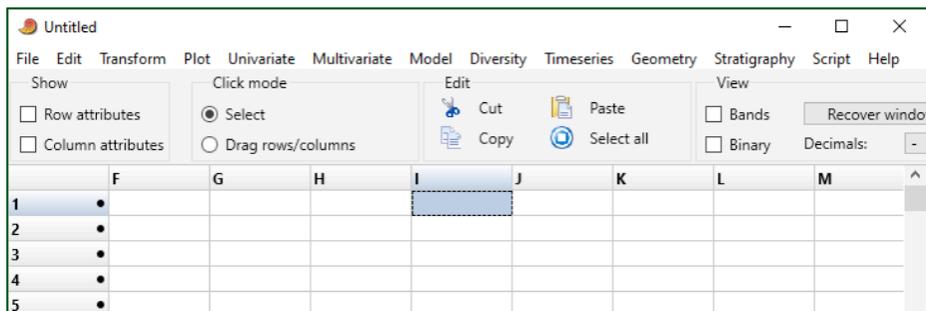



Figura 4.2: Área de trabalho do Past versão 4.06b (2021).

Abas do Past

As abas agrupam as opções e funções que o Past pode desempenhar e variam dependendo da versão. Na versão 4.06b, o Past apresenta as seguintes abas **Figura 4.3.**

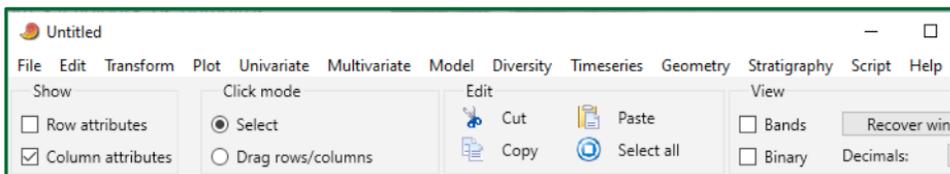



Figura 4.3: Abas do Past versão 4.06b

No menu **File** estão agrupadas as opções, como **New** (novo), **Open** (abrir), **Save** (salvar), além de **Exit** (Sair). No menu **Edit** agrupam funções básicas para a edição dos dados, com **Undo** e **Redo** (desfazer e refazer), **Copy** e **Paste** (copiar e colar), **Select all** (selecionar tudo). Ainda na aba Edit há opções de edição dos dados mais avançados, com **Insert more row** e **Insert more Column** (inserir linhas ou colunas), **Remove e Remove uninformative row/column** (remover e remover não informados) para remover um valor selecionado ou toda linha/coluna, Replace (substituir) para substituir um valor específico em toda a base de dados, entre outras funções. **Figura 4.4.**

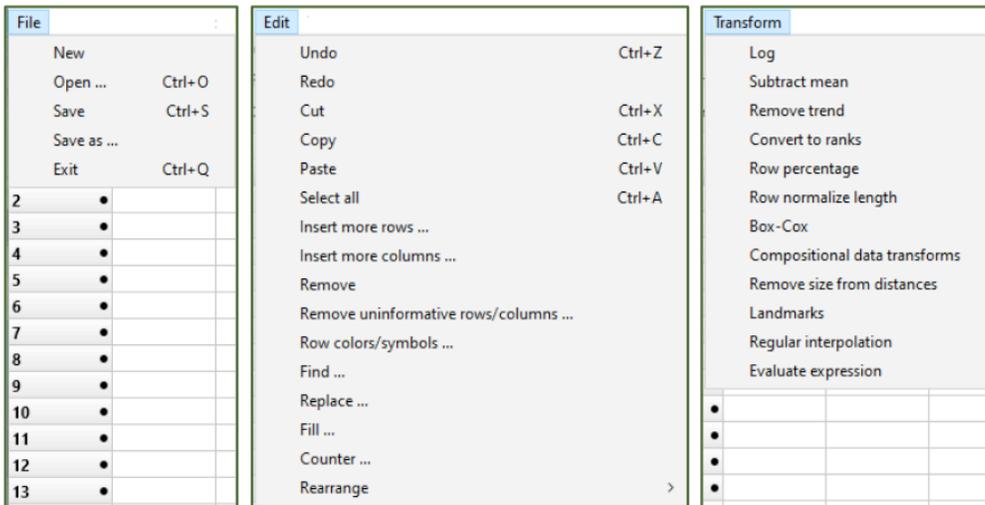



Figura 4.4: Menu File, Edit e Transform do Past versão 4.06b

Há ainda Abas dentro de Abas que agrupam funções semelhantes, "gavetas" facilmente identificadas pelo sinal de MAIOR (**>**), que ao serem clicadas abre outras funções disponíveis no Past. Abaixo na **Figura 4.5** está apresentando no menu **Univariate**, e na aba **Anova etc**, estão agrupados os testes de anova disponíveis no Past.

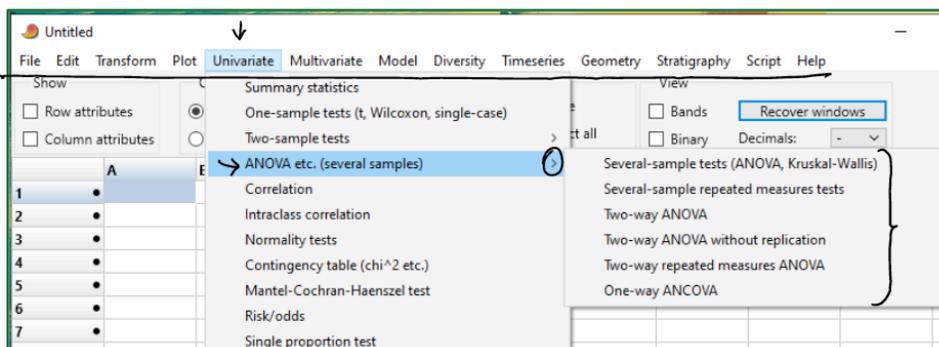



Figura 4.5: Interface de Abas (menu Univariate) e aba (Anova etc.) do Past

Dados no Past

Os dados podem ser inseridos no Past de várias formas diferentes. Desde inserindo os dados manualmente ou copiando de uma planilha do Excel utilizando os atalhos de copiar e colar (Ctrl+C e Ctrl+V). As versões mais recentes do Past ganharam o suporte para ler arquivos no formato ".xls", mesmo formado utilizando dos arquivos do Excel,

possibilitando abrir planilhas do Excel.

Vale ressaltar que o Past interpreta tanto a vírgula (,) quando o ponto (.) como separadores de decimais. E amostras com dados faltando devem ser representados com ponto de interrogação (?).

Para inserir dados de uma tabela que tenha legendas nas colunas no Past primeiro é preciso marcar a opção **Column Attributes** para indicar ao Past que a primeira linha das colunas são as legendas das variáveis. E possível ainda indicar ao Past se os dados pertencentes a coluna são valores numéricos ou tipo grupo (**Group**), e utilizado para indicar grupos distintos na tabela no formato longo. **Figura 4.6.**

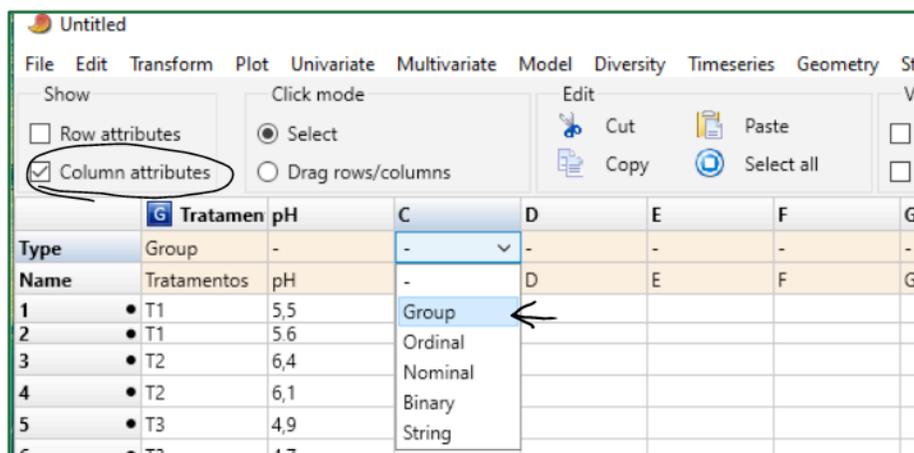



Figura 4.6: Opção Column Attributes Ativa

Para inserir dados de uma tabela formato longo que possuam legendas nas linhas (Mesmo tratamento em diferentes linhas), é preciso marcar a opção **Row Attributes** para indicar ao Past que linhas da coluna são grupos. Para as linhas, pode-se especificar a cor do símbolo que aparecerá nos gráficos (Color), o tipo de símbolo (Symbol) e um nome para cada sujeito, ou unidade amostra (Name). **Figura 4.7.**

Figura 4.7: Opção Row Attributes ativa

Após finalizar a edição dos títulos da tabela, pode-se desmarcar as caixas de Row e Column atributes para que as escolhas e os títulos fiquem gravados.

Figura 4.8**:** Tela inicial após configurar o Row e Column attributes.

PLOTANDO GRÁFICOS NO PAST

O Past tem a capacidade de gerar vários tipos de gráficos, para plotar um gráfico é necessário clicar no menu **Plot**, e selecionar um dos diferentes tipos de gráficos disponíveis.

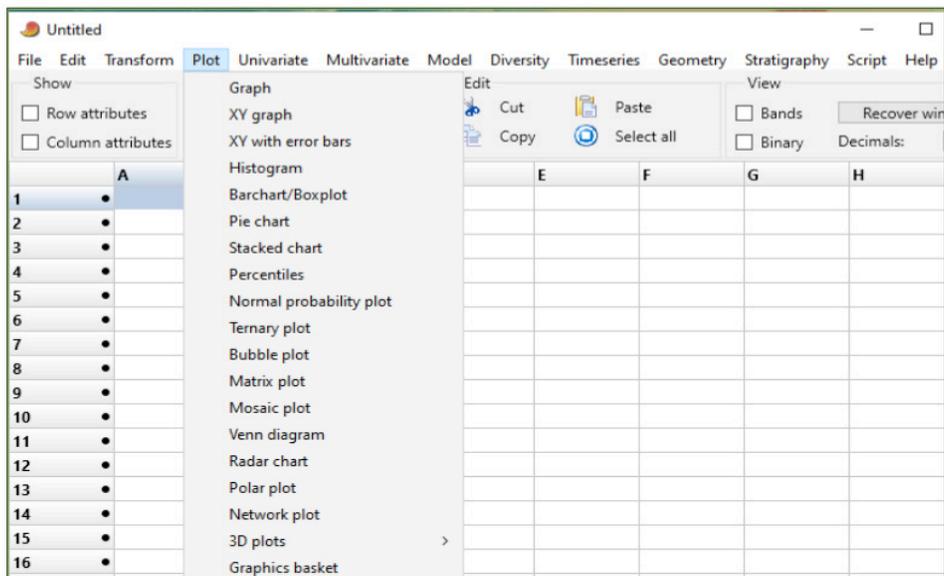



Figura 4.9: Menu Plot e as opções de gráficos disponíveis no Past.

Gráfico de linhas no Past

Para inserir um gráfico de linhas é necessário selecionar as colunas de interesse e clicar em Plot > Graph, e será produzida uma linha para cada coluna.

Figura 4.10**:** Inserindo gráfico de Linhas.

Há como escolher 6 (seis) estilos diferentes de gráficos de linhas na caixa de opção **Plot style**, dos estilos mais comuns, como **Line** (linhas), Points (pontos) ou **Line+Points** (linhas com pontos) entre outros, **Figura 4.11**. Abaixo da caixa de opção Plot style há opção de edição rápida do gráfico como e **Silhouette** (Cor da área abaixo do gráfico), **Log**

Y (log base 10), 3-**point average** (média de 3 pontos), **Row labels** (rótulos de linha), **Flip axes** (virar eixos), **Plot rows** (plotar linhas). Para edições de aparência, há o menu Graph settings, que possui mais ajustes.

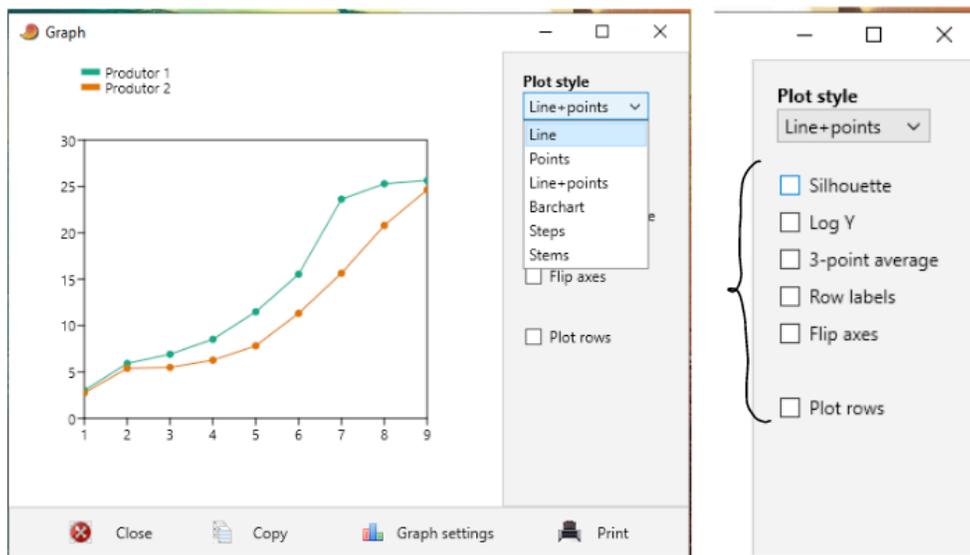



Figura 4.11: Opções de edição de gráficos de linhas.

Gráfico de Dispersão XY no Past

Para inserir um gráfico de dispersão é necessário selecionar as colunas com as variáveis X e Y de interesse, e clicar em **Plot** > **XY graph**. Na caixa de opção **Plot style** estão disponíveis configurações de edição rápida do gráfico de dispersão, como **Points** (pontos), **Line** (linhas) e **Silhouette** (Cor da área abaixo do gráfico) entre outros. Para configurações de aparência do gráfico, há o menu **Graph settings.**

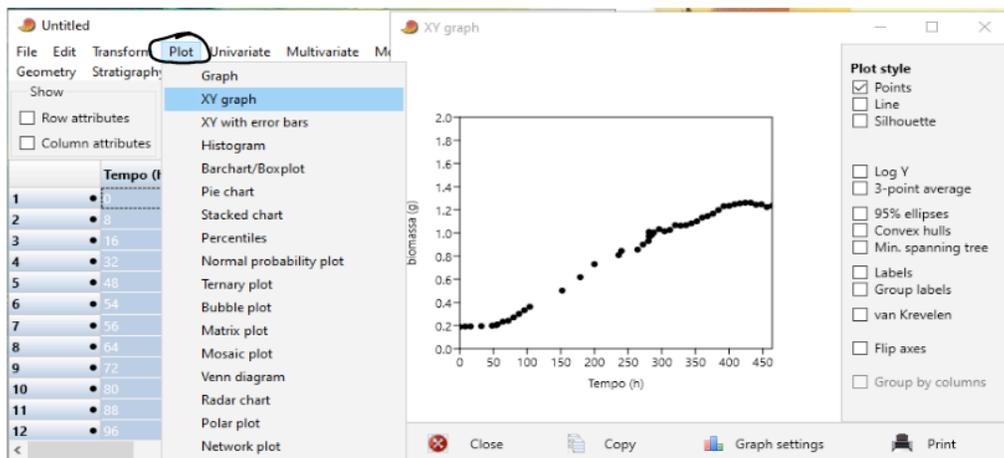



Figura 4.12: Inserindo gráfico de dispersão.

Gráfico de Histograma no Past

Para inserir um gráfico do tipo histograma é necessário selecionar a coluna com as amostras observadas, clicar em **Plot** > **Histogram**. Nas opções de edição rápida do histograma, em **Bins** pode definir o número de classes do histograma, em **Star** e **End** é possível alterar o valor inicial e o final do histograma, a opção **Fit normal** adiciona a linha da normal, **Kernel density** adiciona estimativa de densidade kernel, **Flips** inverte os eixos. Para configurações de aparência do gráfico estão dispostas no menu Graph settings.

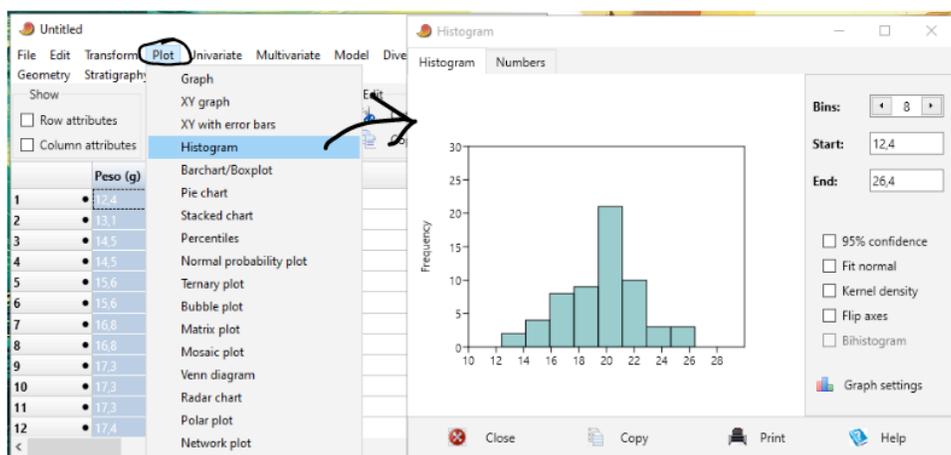



Figura 4.13: Inserindo gráfico de Histograma.

Gráficos de Barras e Boxplot

Para inserir um gráfico de barras ou boxplots os dados podem estar organizados tanto na tabela de formato longo, quanto na tabela de formato largo. Para os dados em formato longo, a coluna com os nomes dos grupos deve estar na especificada como “Group”. Para inserir o gráfico de barras, é necessário selecionar as colunas e clicar em **Plot** > **Barchart/Boxplot**. Na caixa de opção **Plot style** é possível escolher 7 diferentes de gráficos, e logo a baixo estão disponíveis algumas configurações de edição rápida do gráfico.

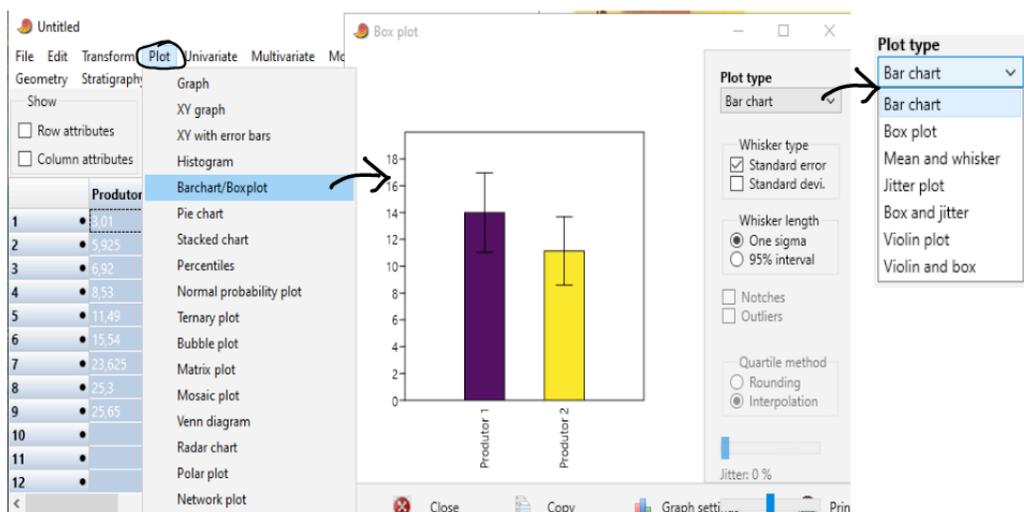



Figura 4.14: Inserindo gráfico de Barras e Boxplot.

Gráficos de Pizza

Para inserir um gráfico de pizza ou setores os dados podem estar organizados tanto na tabela de formato longo, quanto na tabela de formato largo. Para inserir o gráfico de pizza, é necessário selecionar as colunas e clicar em **Plot** > **Pie chart**. Uma nova janela que se abrirá com o gráfico de pizza, logo ao lado estão disponíveis algumas configurações de edição rápida do gráfico.

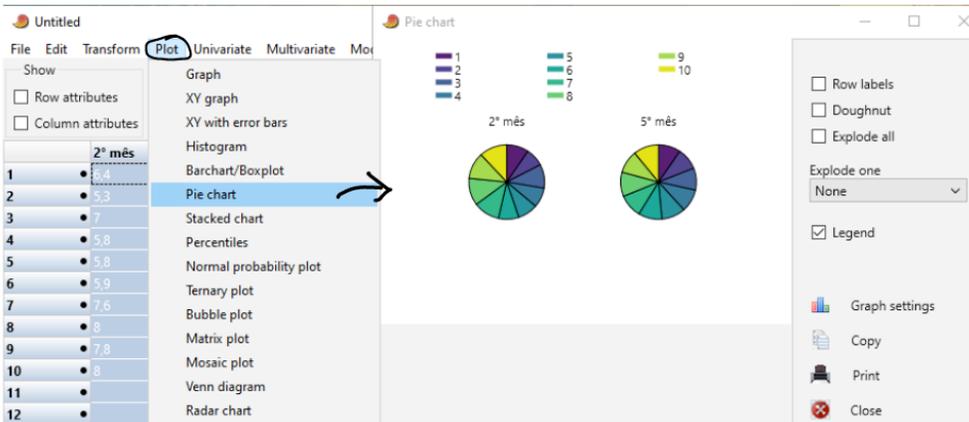



Figura 4.15: Inserindo gráfico de pizza.

ESTATÍSTICA DESCRITIVA

No Past é possível realizar a análise descritiva de forma bastante simples, é necessário selecionar as colunas com os dados de interesse e clicar no menu **Univariate** > **Summary statistics**, onde irá gerar os principais índices de estatística descritiva separadamente para cada coluna.

	Variável X	Variável Y
N	18	18
Min	0.9	1.2
Max	6.9	7.5
Sum	70	71.2
Mean	3.888889	3.955556
Std. error	0.4402086	0.4464253
Variance	3.488105	3.58732
Stand. dev	1.867647	1.894022
Median	4	3.6
25 prcntil	2.175	2.475
75 prcntil	5.425	5.725
Skewness	0.05888355	0.1793562
Kurtosis	-1.028998	-0.9050762

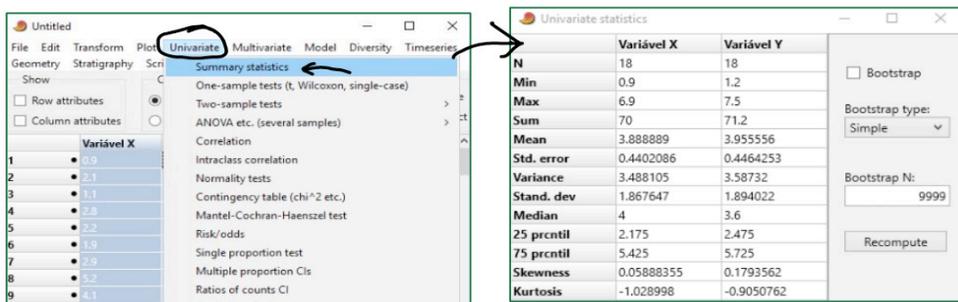



Figura 4.16: Aba Univariate, análise Summary statistics

Alguns dos índices calculados na função Summary statistics são:

- **N:** número de valores utilizados para o cálculo das estatísticas;
- **Min:** valor mínimo;
- **Max:** valor máximo;
- **Sum**: soma;
- **Mean:** média aritmética;

- **Std. Error:** erro padrão amostral, que é igual ao desvio padrão amostral dividido pela raiz quadrada de N;
- **Variance:** variância amostral, que é igual ao desvio padrão amostral ao quadrado;
- **Stand. dev:** desvio padrão amostral;
- **Median:** mediana amostral, correspondendo ao percentil de 50% ou ao 2º quartil;
- **25 prcntil:** percentil de 25%, correspondendo ao 1º quartil;
- **75 prcntil:** percentil de 75%, correspondendo ao 3º quartil;
- **Skewness:** assimetria da distribuição de probabilidade amostral;
- **Geom. Mean:** média geométrica;
- **Coeff. Var:** coeficiente de variação

TESTE T-STUDENT NO PAST

O teste T-Student é um teste estatístico usando na comparação de médias. Este teste tem diversas variações de aplicação, sendo usado na comparação de duas (somente duas) amostras, o teste tem base no teste de hipótese (apresentado no Capítulo 1), hipótese nula (H_0) e hipótese alternativa (H_1), onde o teste T-Student irá testar se a hipótese nula (H_0) é verdadeira ou falsa. A qual por meio do teste t, testa a hipótese que um as médias de duas amostras são iguais ou é igual a um valor fixo previamente estabelecido de acordo com os objetivos da pesquisa.

(4.1): $$t = \frac{\bar{X}}{\sqrt{\frac{s^2}{n}}}$$ Fórmula do t calculado para amostras independentes

O teste t utiliza uma própria distribuição, a Distribuição T, que deriva da distribuição normal. O teste t consiste na comparação de dois índices, o $t_{tabelado}$ e $t_{calculado}$. O $t_{tabelado}$ é um valor pré-determinado encontrado através do nível significância (α) e os graus de liberdade (GL), adotado na pesquisa, consultando na tabela da distribuição t. O $t_{calculado}$ é um valor calculado através de fórmulas matemáticas (**Fórmulas 4.1 e 4.2**). Se o $t_{calculado}$ é maior que o $t_{tabelado}$, rejeitamos H_0, no nível de significância da tabela na qual estava o t que obtivemos.

O teste T-Student tem dois variantes, o teste t para amostras independentes ou não pareadas e o teste t para amostra dependentes ou pareadas. O teste t para amostras

independentes compara as médias de duas amostras não se relacionam, como índices de qualidade de água de 2 viveiros, ou índices zootécnicos de 2 animais ou lotes de animais, comparar o desempenho de tipos diferentes de manejos, comparar produção do “produtor A” com o “produtor B”. Compara ainda a média de uma amostra com uma outra média dada, por exemplo, comparar uma amostra de índices zootécnicos dos peixes do cultivo com os índices ideais na literatura para o cultivo.

O teste t para amostras dependentes compara as médias de duas amostras que se relacionam, os dados das amostras são dependentes, como a comparação das amostras de antes e depois de um determinado tratamento ou manejo em um lote de animais.

(4.2):

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}}$$

Fórmula do t calculado para amostras dependentes

No Past o teste T-Student pode ser realizado com uma amostra, ou com duas amostras univariadas (em duas colunas), nas duas formas são feitos testes paramétricos clássicos onde há comparação das médias e variâncias.

Teste T one-sample

No teste t one-sample (teste t de uma amostra) é usado para investigar se é provável que uma amostra tenha sido retirada de uma população com uma média dada (teórica), a qual pretende comparar com a média da amosta. Para realizar o teste é preciso selecionar a coluna com os dados de interesse, e clicar no menu **Univariate** > **One sample test**. **Figura 4.17.**

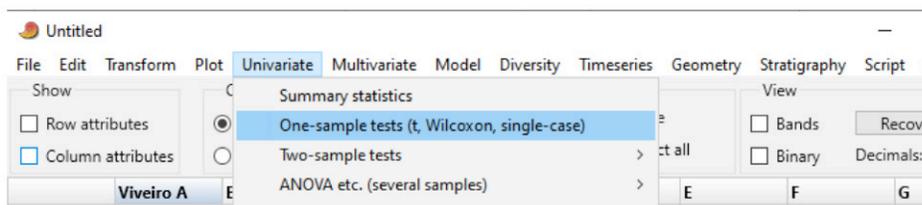



Figura 4.17: Aba Univariate, testes de One-samples tests.

Após clicar na aba **One-sample tests** se abrirá uma nova janela, onde deve ser inserida a média teórica,. A média teórica é inserida na caixa de texto **Given mean** (Média

dada), em seguida é só clicar em **Compute** para gerar o resultado do teste t.

Figura 4.18: Aba Univariate, testes de One-samples tests.

Para interpretar o resultado do teste One-sample t test (teste t de uma amostra) é bem simples, o Past irá exibir os resultados como, média dada (given mean), média da amostra (sample mean), diferença entre as médias(difference), o $t_{calculado}$ (t), o p-Valor (p same mean), além de uma mensagem informando se as médias tem ou não diferença significativa, exibida na parte inferior da janela. **Figura 4.19**.

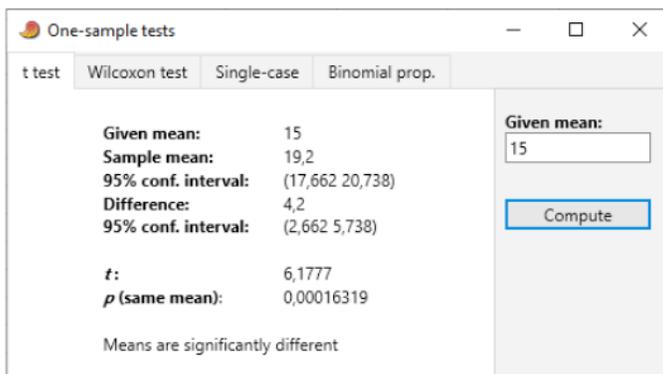



Figura 4.19 resultado d teste T-Student de uma amostra.

Testes two-samples

Test T two-sample (teste t de duas amostras) pode ser realizado de uma maneira simples, para isso as amostras devem estar em tabelas no formato largo, após selecionar as duas colunas com os dados de interesse, e ao clicar no menu **Univariate** > **Two-samples tests**, abrirá três outras opções de testes, **Two-sample tests** (**T, F, Mann-Wh, Kolm-**

Sm etc), (teste de duas amostras), **Two-sample paired tests** (teste de duas amostras pareadas), **F and t tests from parameters** (Testes F e t de parâmetros). **Figura 4.19.**

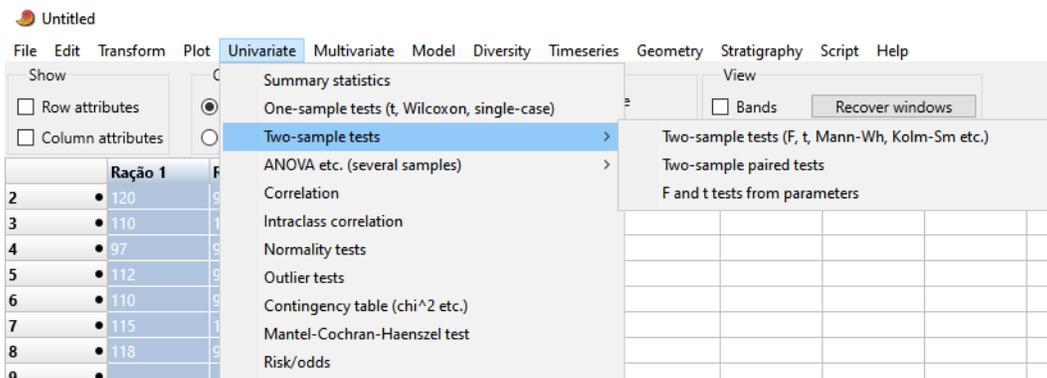



Figura 4.20**:** Aba Univariate, testes de Two-samples tests.

Ao clicar na opção **Two-sample tests (T, F, Mann-Wh, Kolm-Sm etc.)** uma nova janela se abrirá com os resultados de vários testes, entre eles, o teste T, o teste F, teste de Mann-Whitney, Mood median, Kolm-Smirnov, Andreson, Epps-Singleton e Coeff of variation. **Figura 4.21.**

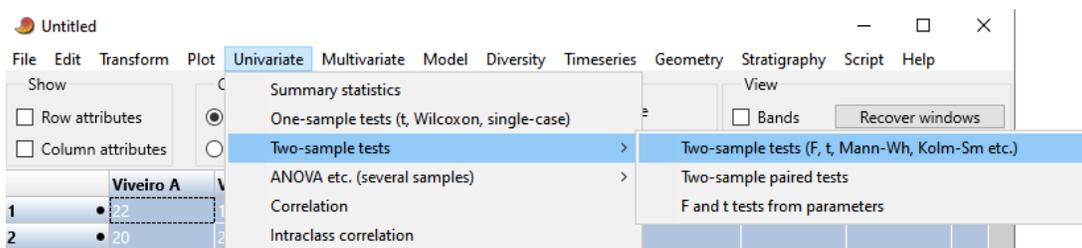



Figura 4.21: Aba Two-samples tests, opção Two-sample tests (T, F, Mann-Wh, Kolm-Sm etc.)

Na janela que se abriu (**Figura 4.22**), interpretar o resultado do teste t é simples, diferentemente da forma convencional, não é necessário consultar o $t_{tabelado}$ uma tabela, e muito menos calcular o $t_{calculado.}$ A interpretação do teste é realizada comparando os valores do p-Valor (**Critical t value**) com o α, o Past adota o valor de α = 0,05. Se o p-Valor for ≤ α, então rejeitasse H_0, mais se o p-Valor for ≥ α, então aceitasse H_0.

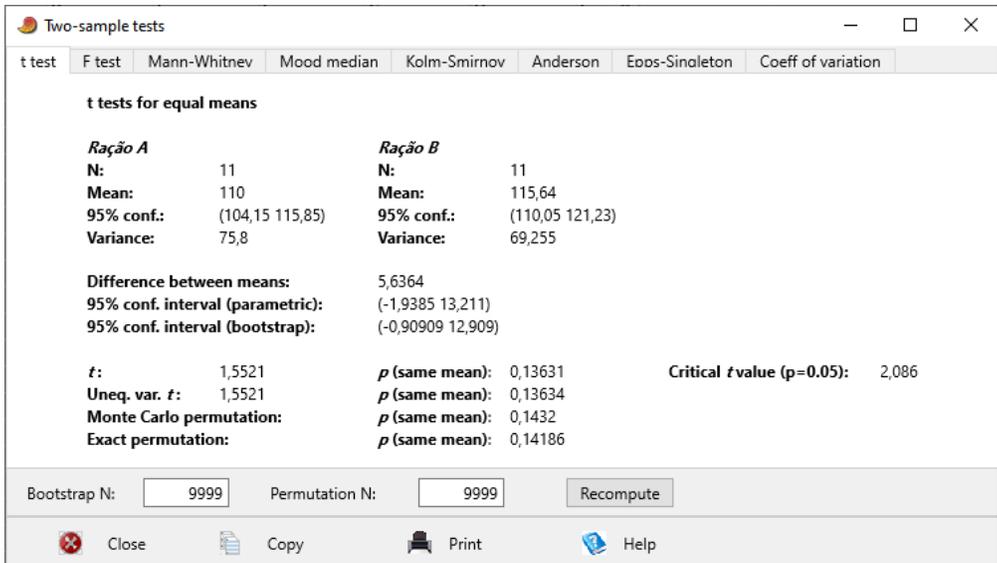



Figura 4.22: Janela Two-samples tests com o resultado do teste T-Student

De maneira simples, se o valor do p-Valor for menor que o valor de α, assumisse que há diferença significativa ao nível de confiança de 95%, então rejeitasse a hipótese nula elaborada no início do experimento.

F test (teste f) ou estatística F deriva da distribuição F. Assim com o teste t para amostras independentes, serve para comparar duas estimativas de variâncias independentes, se assemelhando muito com o teste t, possuindo o $F_{tabelado}$ encontrado na tabela da distribuição f pelo nível de significância (α), e o $F_{calculado}$ calculado pela fórmula. O teste f consiste na comparação dos dois índices $F_{tabelado\ e}$ $F_{calculado}$, se o $_{e}F_{calculado}$ for ≥ $F_{tabelado}$, rejeitasse a hipótese nula, e assumisse que há diferença significativa, mas se o $_{e}F_{calculado}$ for < $F_{tabelado}$, aceitasse a hipótese nula.

No Past o resultado do este f é similar ao teste t, a interpretação do teste é realizada observando se o p-Valor, nesse caso o critical F value (F-Valor) é maior ou menor que o α, o Past adota o valor de α = 0,05. Se o F-Valor for ≤ α, Então Rejeitasse H_0, mais se o F-Valor for ≥ α, então aceitasse H_0. Assumindo-se que ao nível de confiança de 95%, que há ou não diferença significativa, rejeitasse ou aceitando a hipótese nula elaborada no início do experimento. **Figura 4.23.**

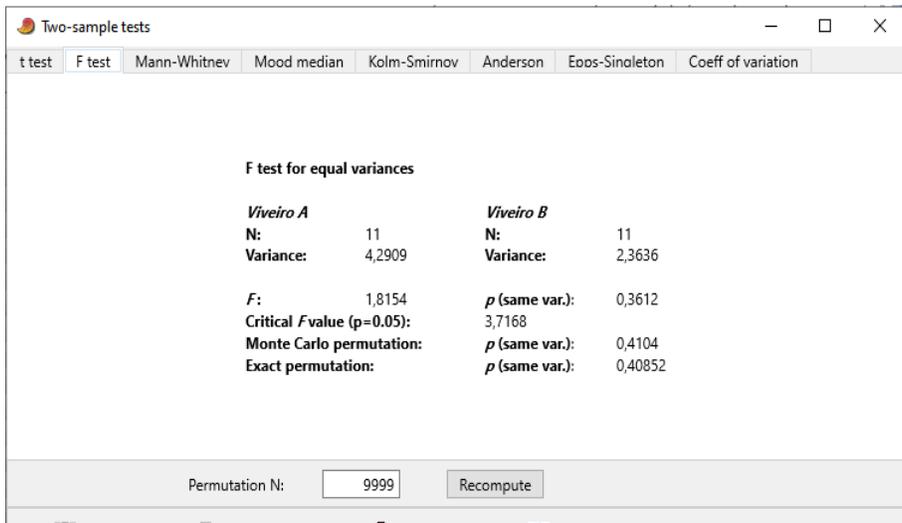



Figura 4.23: Janela Two-samples tests com o resultado do teste F

Mann-Whitney test (Teste de Mann-Whitney) O teste bicaudal U de Mann-Whitney pode ser usado para testar se as medianas de duas amostras independentes são diferentes. Ele não assume distribuição normal, mas assume que as distribuições dos dois grupos têm a mesma forma.

Kolmogorov-Smirnov é um teste não paramétrico que testa se duas distribuições univariadas apresentam a mesma distribuição geral. Em outras palavras, este teste não testa especificamente a igualdade de média, variância ou qualquer outro parâmetro.

Coefficient of variation (Coeficiente de variação), este módulo testa se duas amostras, fornecidas em duas colunas, o coeficiente de variação (ou variação relativa) é definido como a razão do desvio padrão e da média em porcentagem.

Two-sample paired tests (testes de duas amostras pareadas), Três testes estatísticos (um paramétrico, dois não-paramétricos) para duas amostras (colunas) de dados. Os pontos de dados são pareados, significando que os dois valores de cada linha são associados. Para realizar o teste, é necessário clicar no menu **Univariate** > **Two-samples tests** > **Two-sample paired tests**. **Figura 4.24.**

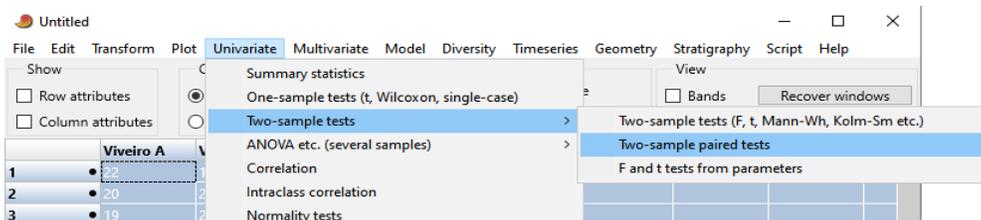



Figura 4.24: Aba Two-samples tests, opção Two-sample paired tests

Ao clicar na opção **Two-sample paired tests** uma nova janela se abrirá com o resultado dos Três testes estatísticos, o t test (teste t para amostras dependentes), o Sign test (teste de sinal), o Wilcoxon test (teste Wilcoxon). **Figura 4.25**

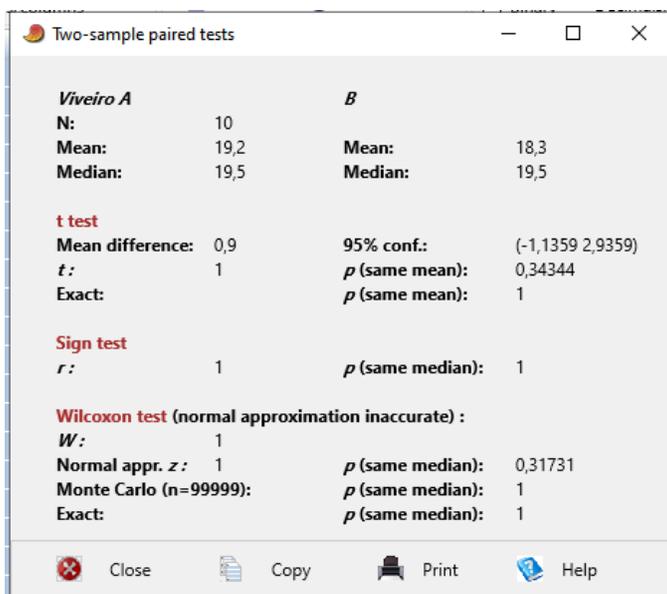



Figura 4.25: Janela com os resultados dos testes da opção Two-sample paired tests

F and t tests from parameters (Testes F e t a partir de parâmetros) útil para realizar analises quando as publicações não fornecem os dados, mas fornecem valores para tamanhos amostrais, média e variância de duas amostras. Estes podem ser inseridos manualmente usando a opção “F and t from parameters” no menu. Esse módulo não usa dados da planilha.

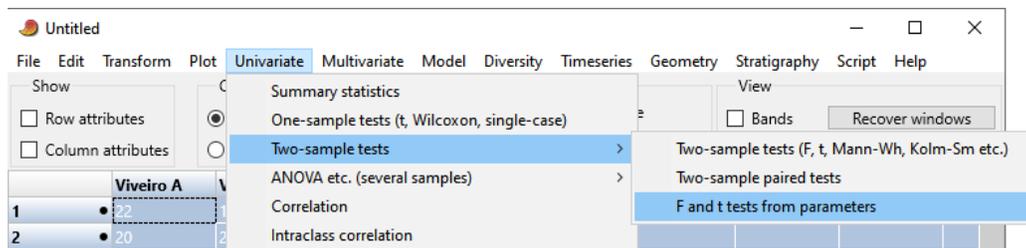



Figura 4.26: Aba Two-samples tests, opção F and t from parameters

ANÁLISE DE VARIÂNCIA ANOVA

A ANOVA análise de variância é um procedimento estatístico que compara as variações dentro de cada tratamento e entre os tratamentos adotados no experimento. No teste de ANOVA o tipo de delineamento experimental adotado no experimento é muito importante, pois o delineamento irá determinar números de fatores, de tratamentos e de repetições que o experimento irá adotar.

One-way ANOVA no Past

A One-Way ANOVA ou unifatorial normalmente é usada para testar diferenças entre pelo menos três grupos, uma vez que a comparação entre dois grupos pode ser obtida através do teste t.

Um experimento com um fator (X) com três tratamentos (T1,T2,T3) e uma variável respota contínua (Y). A One-Way ANOVA verificará se os tratamentos possuem efeitos diferentes na variável resposta (Y). Importante lembrar: a Hipótese nula diz que as médias são iguais entre os grupos, ou seja, as médias entre T1, T2 e T3 devem ser iguais, não existe diferença entre os tratamentos.

ANOVA (análise de variância) unifatorial é um para testar a hipótese nula de que uma série de amostras univariadas (em colunas) são tomadas de populações com a mesma média. Caso a amostra não derive de uma distribuição normal, o teste não-paramétrico de Kruskal-Wallis deve ser usado ao invés da ANOVA.

One-way ANOVA DIC no Past

Para realizar a ANOVA One-wey delineamento inteiramente casualizado no Past, os dados devem ser organizados, com os tratamentos nas colunas e as repetições nas linhas.

Exemplo 1 DIC: No intuito de avaliar o potencial de 4 rações na influência do crescimento, em comprimento (cm) do tambaqui cultivado em 120 dias, obteve-se os resultados: **Tabela 4.1**

Repetições	Ração A	Ração B	Ração C	Ração D
R1	13,39	17,1	22,15	25,5
R2	8,2	18,25	22,79	24,8
R3	13,61	18,78	22,76	24,4
R4	12,74	19,76	24,01	26,3

Tabela 4.1: Médias de comprimento (cm) de tambaqui de cultivo alimentos com 4 rações diferentes.

A **Tabela 4.1** apresenta os dados obtidos de um experimento, que investiga os efeitos no crescimento (cm) de tambaquis alimentos com 4 rações diferentes. Para isso experimento idealizado seguindo o delineamento inteiramente casualizado, com 4 tratamentos (raçãos A, B, C e D), e 4 repetições (R1, R2, R3 e R4) dispostos em 16 tanques de engorda. Na **Tabela 4.1** estão as médias do comprimento (cm) de 10 tambaquis, para cada repetição dos 4 tipos de ração, totalizando 160 tambaquis, 40 animais para cada tipo de ração testado, divididos em 16 tanques.

Os dados da **Tabela 4.1** já estão organizados para realizar a ANOVA One-wey, para isso é necessário importar copiando a planilha de dados do Exce para o Past, como os dados tem legenda nas linhas (repetições) e nas colunas (rações), antes colar os dados no Past, as opções **Row attributes** e **Column attributes** devem ser marcadas, para que o Past identifique os tratamentos e as repetições. **Figura 4.27.**

Untitled

File Edit Transform Plot Univariate Multivariate Model Diversity Timeseries Geometry Stratigraphy Script Help

Show: Row attributes, Column attributes | Click mode: Select, Drag rows/columns | Edit: Cut, Copy, Paste, Select all | View: Bands, Binary, Recover w, Decimals:

	Color	Symbol	Name	Ração A	Ração B	Ração C	Ração D
Type			-	-	-	-	-
Name	Black	Dot	Repetições	Ração A	Ração B	Ração C	Ração D
R1	Black	Dot	R1	13,39	17,1	22,15	25,5
R2	Black	Dot	R2	8,2	18,25	22,79	24,8
R3	Black	Dot	R3	13,61	18,78	22,76	24,4
R4	Black	Dot	R4	12,74	19,76	24,01	26,3

Figura 4.27: Inserindo dados com legendas nas linhas e colunas com opções Row attributes e Column attributes ativa

	Ração A	Ração B	Ração C	Ração D	E	F	G
R1	13,39	17,1	22,15	25,5			
R2	8,2	18,25	22,79	24,8			
R3	13,61	18,78	22,76	24,4			
R4	12,74	19,76	24,01	26,3			

Figura 4.28: Dados da Tabela 4.1 inseridos no Past

Após a inserção dos dados da **Tabela 4.1** no Past, é preciso desmarcar as opções **Row attributes** e **Column attributes**, e após selecionar todos os dados, é necessário clicar no menu **Univariate** > Aba **ANOVA etc (several samples)** > opção **Several-sample tests (ANOVA, Kruskai=Wallis). Figura 4.29.**

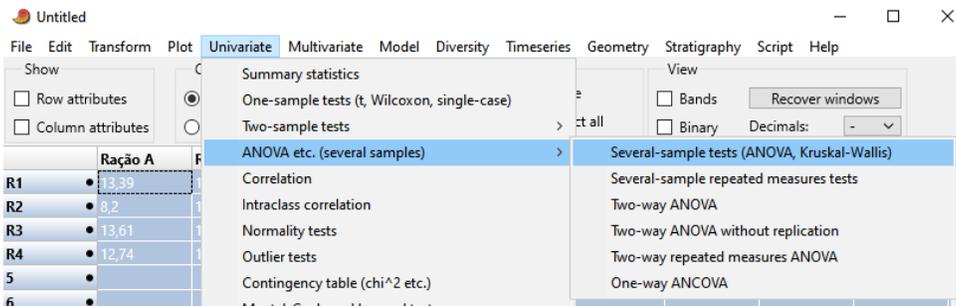



Figura 4.29: menu Univariate > Aba ANOVA etc (several samples) > opção Several-sample tests (ANOVA, Kruskai=Wallis)

Após clicar na opção **Several-sample tests (ANOVA, Kruskai=Wallis)** uma nova janela se abrirá com os resultados da ANOVA One-way. Entre os resultados da ANOVA estão os valores das médias entre os grupos e dentro dos grupos (Between groups e Within groups), entre outros índices calculados na ANOVA. Para a interpretação do resultado ANOVA, é necessário observar o valor do **p(same)** ou p-Valor, se o p-Valor for ≤ α, então rejeitasse H_0 e assumisse que há diferença significativa.

Figura 4.30: Resultado da ANOVA One-way dos dados da Tabela 4.1.

O p-Valor da ANOVA One-way da **Tabela 4.1**, foi um valor muito pequeno em forma de notação cientifica (1,647E-07), se valor é bem menor que o valor do α (0,05), então assumisse que as 4 rações são diferença entre si. A ANOVA somente diz se há ou não diferença, mas não aponta qual a diferença.

Para saber qual a diferença entre os tratamentos é utilizado o teste de Tukey's pairwise, parando par-a-par os tratamentos. **Figura 4.31**

	Ração A	Ração B	Ração C	Ração D
Ração A		0,000268	1,372E-06	1,644E-07
Ração B	8,632		0,005929	0,0001788
Ração C	14,56	5,928		0,1825
Ração D	17,65	9,018	3,09	

Figura 4.31: Resultado do teste de Tukey's pairwise dos dados da Tabela 4.1

Exemplo 2 DIC: Para estudar o efeito da estocagem no crescimento dos peixes, alevinos de tilápia com peso médio de 18,85g foram estocados em gaiolas de 1,7 m³ cada, nas seguintes densidades:

- T75= densidade de 75 peixe/m³

- T150= densidade de 150 peixe/m³
- T300= densidade de 300 peixe/m³
- T600= densidade de 600 peixe/m³

Em um delineamento inteiramente casualizado, (4 tratamentos e 6 repetições), cultivados por 196 dias, das obtido estão dispostos na **Tabela 4.2**

			Tanques		
Tratamentos	1	2	4	5	**6**
T75	91,2	91,3	93,1	89,2	96,8
T150	84,1	84,5	84,9	83,1	83,2
T300	86,9	86,2	87,5	88,3	92,1
T600	77,3	76,1	78,1	81,8	79,6

Tabela 4.2: Médias de peso (g) de alevinos de tilápia cultivados em diferentes estocagens durante 196 dias de cultivo

Como a **Tabela 4.2** está organizada com os tratamentos nas linhas e as repetições nas colunas, primeiro é necessário inverter a tabela para realizar ANOVA One-way. É possível interver os dados tanto no Excel, com a função **Transpor**, quanto no próprio Past com a função Tranpose.

Para transpor a tabele no Excel, primeiro deve copiar a tabela, utilizando o Ctrl+C, clicar em uma célula vazia, pressionar o botão direito do mouse, clicar em Colar especial, e uma nova janela se abrirá, nela é só marcar a opção Tranpor, e clicar no OK.

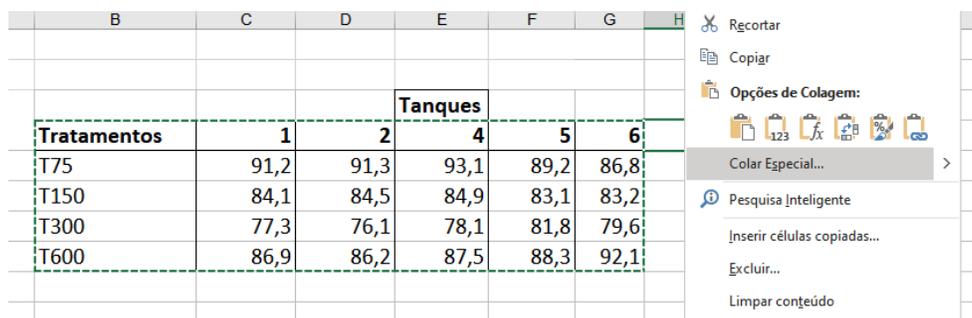



Figura 4.32: Colar Especial do Excel.

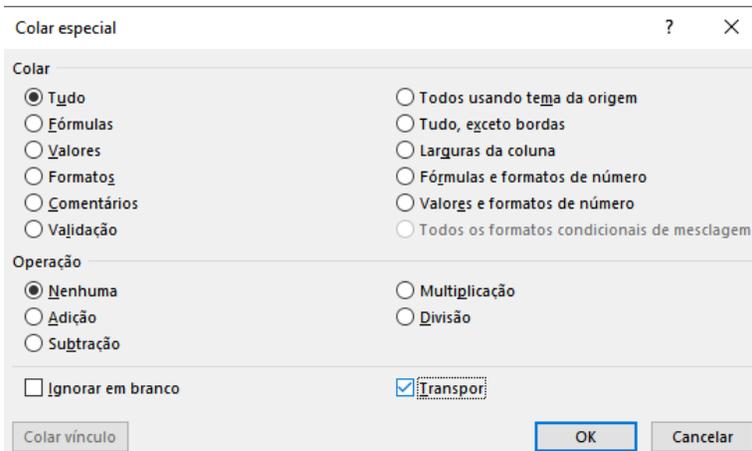



Figura 4.33: Opção transpor do Excel.

Após clicar no OK o Excel irá transpor a tabela, resultando em uma organização apresentada na **Figura 4.34**. As legendas “Taques e Tratamentos” podem ser excluídas, por não serem necessárias no Past, após isso a tabela pode ser editada no Past.

	Tratamentos	T75	T150	T300	T600
	1	91,2	84,1	77,3	86,9
	2	91,3	84,5	76,1	86,2
Tanques	**4**	93,1	84,9	78,1	87,5
	5	89,2	83,1	81,8	88,3
	6	86,8	83,2	79,6	92,1

Figura 4.34: Resultado da função transpor do Excel.

Para inverter o a **Tabela 4.2** no Past, é necessário copiar os valores para Past atribuindo legendas nas linhas e colunas (Row attributes e Column attributes), com tudo selecionado e só clicar no menu **Edit** > aba **Rearrange** > função **Transpose. Figura 4.34.**

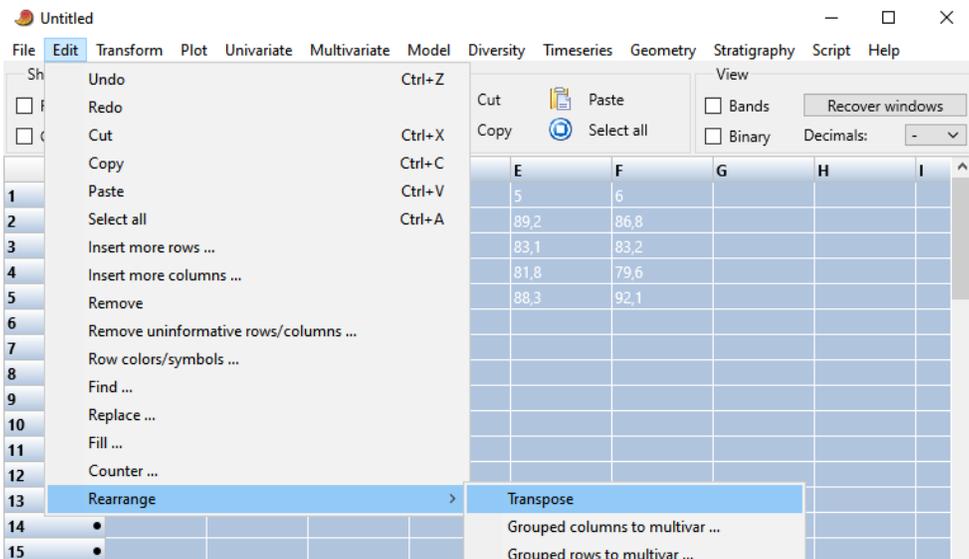



Figura 4.34: Menu Edit, Função Transpose do Past.

Após função Transpose os dados da **Tabela 4.2** iram ficar semelhante ao apresentado na **Figura 4.35**.

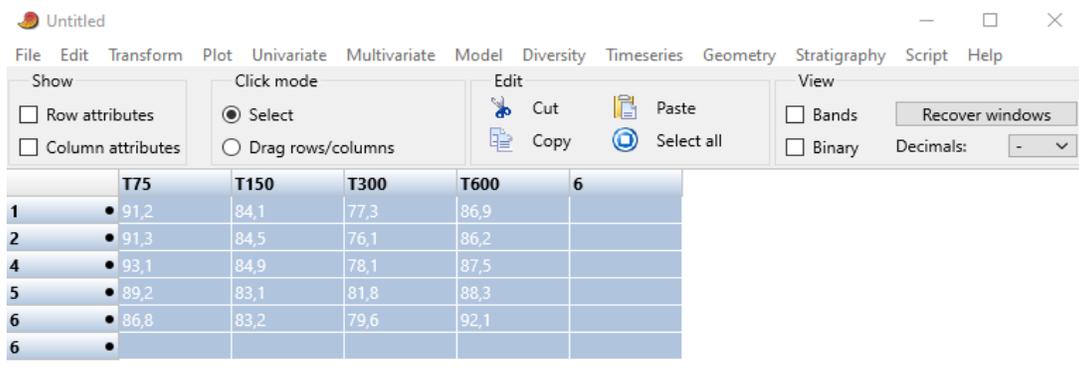



Figura 4.35: Resultado da função Transpose do Past.

Os dados apresentados na **Figura 4.35** já estão organizados para realizar a ANOVA One-way, é só clicar no menu **Univariate** > Aba **ANOVA etc (several samples)** > opção **Several-sample tests (ANOVA, Kruskai=Wallis).**

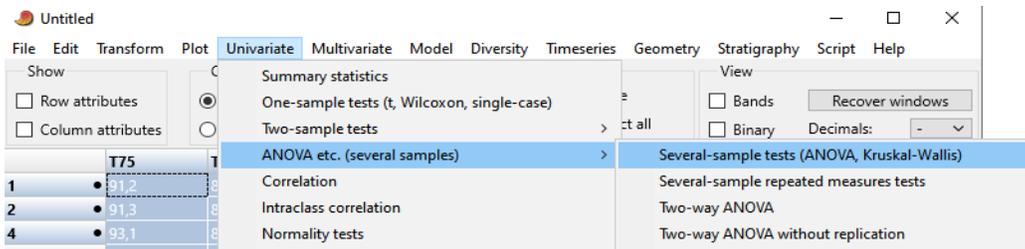



Figura 4.36: menu Univariate > Aba ANOVA etc (several samples) > opção Several-sample tests (ANOVA, Kruskai=Wallis).

Após clicar na opção **Several-sample tests (ANOVA, Kruskai=Wallis)** uma nova janela se abrirá com os resultados da ANOVA One-way. **Figura 4.37.**

Test for equal means

	Sum of sqrs	df	Mean square	*F*	*p* (same)
Between groups:	402,798	3	134,266	32,33	5,046E-07
Within groups:	66,448	16	4,153		**Permutation *p* (n=99999)**
Total:	469,246	19			1E-05

Components of variance (only for random effects):
Var(group): 26,0226 **Var(error):** 4,153 **ICC:** 0,862372

***omega2*:** 0,8245

Levene´s test for homogeneity of variance, from means *p* (same): 0,3445
Levene´s test, from medians *p* (same): 0,6535

Welch *F* test in the case of unequal variances: *F*=22,82, df=7,79, *p*=0,0003229

Figura 4.37: Resultado da ANOVA One-way dos dados da Tabela 4.2

Como o valor do **p(same)** é muito pequeno (5,046E-07), menor que o valor do α (0,05), assumisse que há diferença significativa para as diferentes densidades de estocagens de alevinos de tilápia.

One-way ANOVA DBC no Past

Para realizar a ANOVA One-wey delineamento em blocos casualizados no Past, diferente da ANOVA DIC, os dados devem ser organizados, com os tratamentos nas linhas e os blocos nas colunas.

Exemplo 1 DBC: Ao testar a influência das rações A, B, C e D na conversão alimentar de alevinos curimatã cultivado durante 2 meses em tanques-rede de 2 m², à densidade de 5 ind/m².

- Ração A = tipo 1 (20% de Proteína)
- Ração B = tipo 1 (30% de Proteína)

- Ração C = tipo 2 (20% de Proteína)
- Ração D = tipo 2 (30% de Proteína)

Foram utilizados viveiros experimentais, e os resultados do experimento estão apresentados na T**abela 4.3**

	Bloco		
Tratamento	**1**	**2**	**3**
A	2,7	2,7	2,6
B	2,4	2,5	2,5
C	2,8	2,7	2,8
D	2,6	2,7	2,6

Tabela 4.3: Médias de peso (g) cultivo alevinos curimatã alimentado com 4 diferentes rações

A **Tabela 4.3** apresenta os dados obtidos de um experimento, que investiga os efeitos no no ganho de peso em alevinos curimatã alimentos com 4 rações diferentes. Para isso experimento idealizado seguindo o delineamento Blocos casualizados, com 4 tratamentos (raçãos A, B, C e D), e 3 blocos (blocos 1, 2 e 3).

Os dados da **Tabela 4.3** já estão organizados para realizar a ANOVA One-wey, para isso é necessário importar copiando a planilha de dados do Exce para o Past, com as opções **Row attributes** e **Column attributes** ativas, e após inserindo os valores devem ser desativadas.

Após isso a ANOVA DBC é similar a ANOVA DIC, é só clicar no menu **Univariate** > Aba **ANOVA etc (several samples)** > opção **Several-sample tests (ANOVA, Kruskai=Wallis).** O resultado da ANOVA DBC está apresentado no Figura 4.38

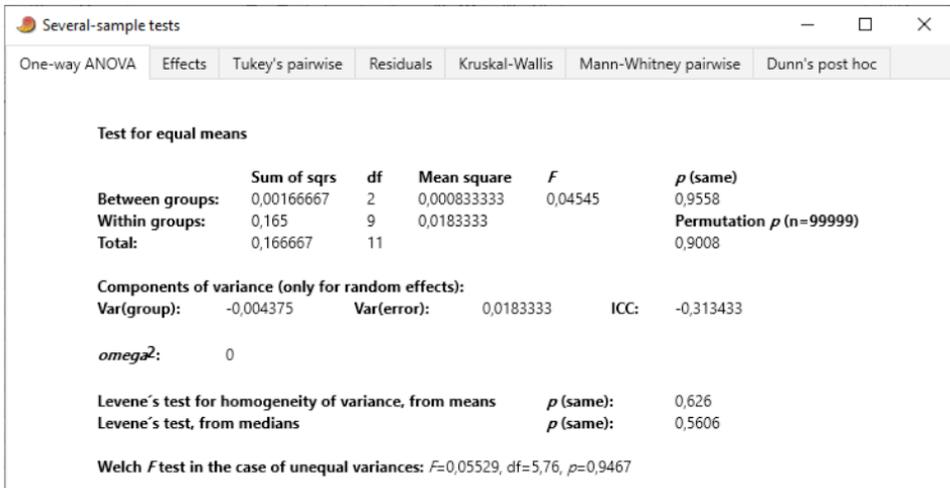



Figura 4.38: Resultado da ANOVA One-way dos dados da Tabela 4.3

Como exibido na Figura 4.37, o valor do p(same) é maior que o valor do α, p(same) = 0,9558 e α = 0,05. Assumisse que não há diferença entre significativa entre as rações.

Two-way ANOVA

A ANOVA Two-way (análise de variância) bifatorial é uma medida estatística para testar a hipótese nula de que uma série de amostras univariadas têm a mesma média em relação a cada um de dois fatores, e que há dependências (interações) entre fatores. Três colunas são necessárias. Primeiro, uma coluna com os níveis do primeiro fator (codificadas como 1, 2, 3 etc), depois uma coluna com os níveis do segundo fator, e finalmente uma coluna com as medidas dos valores correspondentes.

Two-way ANOVA no Past

Para realizar a ANOVA Two-way no Past a planilha dos dados deve estar organizada no formato longo (Apresentado no capítulo 3), com os 2 fatores e as medidas dos valores correspondentes, disposta nas colunas.

Exemplo 3 DIC: Em um delineamento inteiramente casualizado, em esquema fatorial, com 4 repetições, avaliou-se a influência das densidades 20, 40 e 60 tambaquis/m², com peixes de 1 a 4 meses de idade observando os seguintes resultados em peso (g). **Tabela 4.4**.

		Repetições			
Densidade	**Idade**	**1**	**2**	**3**	**4**
D20	**1**	39	33	33	32
D20	**2**	128	115	126	130
D20	**3**	56	210	225	230
D20	**4**	334	324	301	309
D40	**1**	36	24	30	39
D40	**2**	124	136	120	107
D40	**3**	153	219	186	210
D40	**4**	384	381	278	369
D60	**1**	36	34	47	39
D60	**2**	96	115	134	172
D60	**3**	217	251	233	302
D60	**4**	323	338	327	395

Tabela 4.4: Médias de peso (g) de tambaquis de diferentes idades e densidades de cultivado.

Para realizar a ANOVA Two-way com os dados da Tabela 4.4, deve estar organizado no formato longo, semelhante ao apresentado na Tabela 4.5.

Densidaddes	Idade	Peso (g)
D20	I1	39
D20	I1	33
D20	I1	33
D20	I1	32
D20	I2	128
D20	I2	115
D20	I2	126
D20	I2	130
D20	I3	256
D20	I3	210
D20	I3	225
D20	I3	230
D20	I4	334
D20	I4	324
D20	I4	301
D20	I4	309
D40	I1	36
D40	I1	24
D40	I1	30

D40	I1	39
D40	I2	124
D40	I2	136
D40	I2	120
D40	I2	107
D40	I3	153
D40	I3	219
D40	I3	186
D40	I3	210
D40	I4	384
D40	I4	381
D40	I4	278
D40	I4	369
D60	I1	36
D60	I1	34
D60	I1	47
D60	I1	39
D60	I2	96
D60	I2	115
D60	I2	134
D60	I2	172
D60	I3	217
D60	I3	251
D60	I3	233
D60	I3	302
D60	I4	323
D60	I4	338
D60	I4	327
D60	I4	395

Tabela 4.5: Dados da Tabela 4.4 organizados no formato longo.

Para organizar os dados da **Tabela 4.4** no formato longo, semelhante a **Tabela 4.5**, é realizado no Excel, mas não há uma função específica, sendo necessário reorganizar os dados manualmente. Após finalizar a reorganização, os dados da **Tabela 4.5** podendo ser inseridos no Past, atribuindo as legendas das colunas (Column attributes), os 2 fatores (densidade e Idade) devem ser especificados com grupos (group).

Posteriormente para realiza a ANOVA Two-way, é só clicar no menu **Univariate** > Aba **ANOVA etc (several samples)** > opção Two-way ANOVA. **Figura 4.38.**

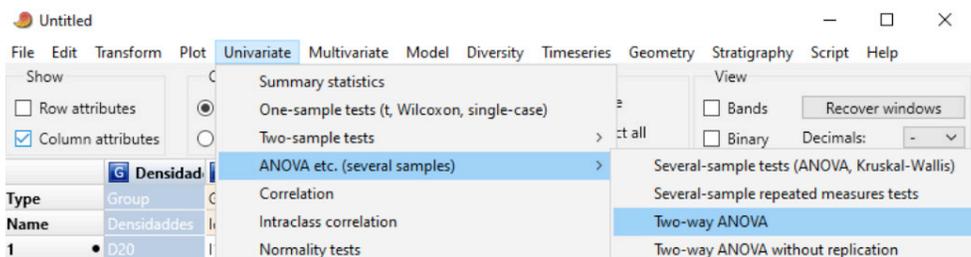



Figura 4.39: Função ANOVA Two-way

Os resultados da ANOVA Two-way está apresentado na figura 4.40, a interpretação do resultado é similar a ANOVA One-way, se o valor do p(same) for menor que o valor de α (0,05), rejeitasse H_0 e assumisse que há diferença significativa.

FIXED-EFFECTS TWO-WAY ANOVA

	Sum of sqrs	df	Mean square	F	p (same)
Densidaddes:	2599,29	2	1299,65	1,998	0,1503
Idade:	613027	3	204342	314,2	7,314E-26
Interaction:	7624,21	6	1270,7	1,954	0,09849
Within:	23414	36	650,389		
Total:	646665	47			

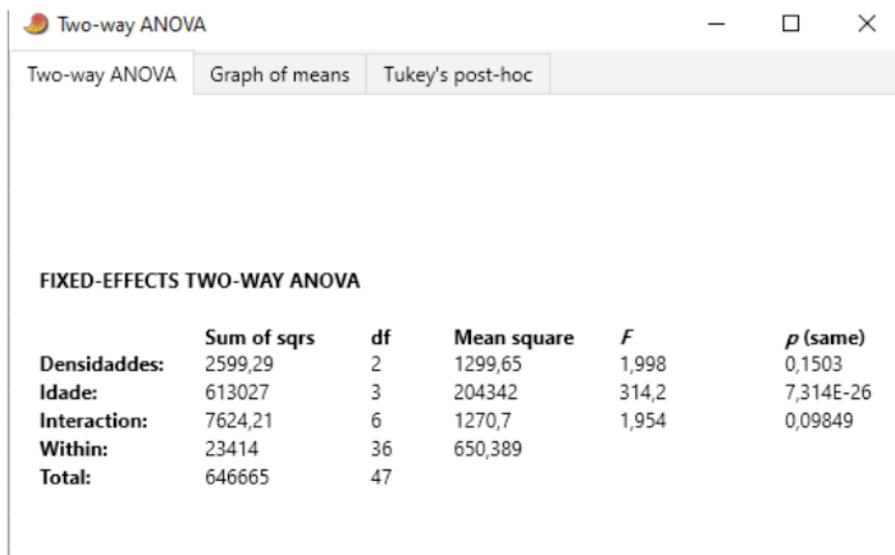



Figura 4.40: Resultado da ANOVA Two-way com os dados da Tabela 4.5.

Como valor de p(same) para a ANOVA two-way com os dados da **Tabela 4.5**, é maior que 0,05, p(same) =0,1503, aceitasse a H_0 e assumisse que no Exemplo 3, não houve diferença significativa.

REFERÊNCIAS

CENTENO, A.J. **Curso de Estatística Aplicada à Biologia**. Goiânia: Editor UFG. 2a Edição. 1999.

DANTAS NEDER, H. **CURSO DE ESTATÍSTICA APLICADA**. Uberlândia, Minas Gerais. v.1. 2015.

DE, W.; BUSSAB, O.; MORETTIN, P. A. **Estatística Básica**. 6. ed. São Paulo: Saraiva, 2010. Disponível em: www.saraivauni.com.br

DOWNING, D.; CLARK, J. **Estatística Aplicada**. Saraiva. 2002.

FERREIRA, P. V. **Estatística experimental aplicada às ciências agrárias.** Ed. UFV. ISBN 978-85-7269-566-4. 2018. 588p.

HAMMER, Ø. **PAleontological STatistics Version 4.07 Reference manual:** Vol. 4.07b (University of Oslo, Ed.; Version 4.07b). Natural History Museum. 2021.

MARTINS, G. A. **Estatística Geral e Aplicada**. 3. ed. São Paulo: Editora Atlas, 2010. v. Único

MENDES, P.P. **Estatística Aplicada à Aquicultura**. Recife: Editora Bagaço. v1, 1999.

ROCHA, A. L. S. **Estatística Aplicada a Engenharia I**. Universidade Federal Do Rio Grande Do Norte Centro De Ciências Exatas E Da Terra Departamento De Estatística, Ed:1. 2011.

SANTOS, E. P. **Dinâmica de populações aplicada à pesca e piscicultura**. São Paulo: Hucitec, Ed. da Universidade de São Paulo. 1978.

VIEIRA, Sonia. **Introdução à bioestatística.** 4. ed. Rio de Janeiro: Elsevier, 2008.

SOBRE OS AUTORES

SUZETE ROBERTA DA SILVA - Professora Adjunta da Universidade Federal do Oeste do Pará, Campus Monte Alegre. Possui graduação (2010), Mestrado (2013), e Doutorado (2016) em Engenharia de Pesca pela Universidade Federal do Ceará. Tem experiência nas áreas de qualidade de água aplicada a aquicultura e biotecnologia. Coordenadora do projeto integrado "Desenvolvimento de alimento alternativo para peixes com resíduos agroindustriais e educação continuada em ciências agrárias em escolas de ensino médio, na UFOPA e produtores rurais" que viabilizou a produção deste material de estudo.

PAULO BRASIL - Possui graduação em Engenharia de Pesca pela Universidade Federal Rural da Amazônia (2011), mestrado em Aquicultura e Recursos Aquáticos Tropicais pela Universidade Federal Rural da Amazônia (2013) e Doutor em Ciências Ambientais do Programa de Pós-Graduação Sociedade, Natureza e Desenvolvimento da Universidade Federal do Oeste do Pará (2019). Atualmente é professor adjunto do Campus Monte Alegre da Universidade Federal do Oeste do Pará. Tem experiência na área de Recursos Pesqueiros e Engenharia de Pesca, com ênfase em avaliação de recursos pesqueiros, ecologia pesqueira, análise de dados, geoprocessamento e geoestatística. Colaborador no projeto integrado "Desenvolvimento de alimento alternativo para peixes com resíduos agroindustriais e educação continuada em ciências agrárias em escolas de ensino médio, na UFOPA e produtores rurais" Desmistificação da estatística nas ciências agrárias".

LUAN PATRICK - Estudante do curso de graduação de Engenheiro de Aquicultura na Universidade Federal do Oeste do Pará, Campus Monte Alegre. Bolsista PEEx no projeto integrado "Desenvolvimento de alimento alternativo para peixes com resíduos agroindustriais e educação continuada em ciências agrárias em escolas de ensino médio" atuando na Desmistificação da estatística nas ciências agrárias, orientado pelo Prof. Paulo Roberto Brasil Santos e Suzete Roberta da Silva.

Introdução à Estatística

Aplicada à aquicultura

Luan Patrick Moura de Souza
Discente de Engenharia de Aquicultura – UFOPA
Bolsista Programa Integrado de Ensino, Pesquisa e Extensão – PEEX/Ufopa

Paulo Roberto Brasil Santos
Doutor em Ciências Ambientais – UFOPA
Docente da UFOPA – Campus Monte Alegre, em Engenharia de Aquicultura

Suzete Roberta da Silva
Doutora em Engenharia de Pesca - UFC
Docente da UFOPA – Campus Monte Alegre, em Engenharia de Aquicultura

Realização:

Apoio:

Ano 2023

Introdução à Estatística

Aplicada à aquicultura

Luan Patrick Moura de Souza
Discente de Engenharia de Aquicultura – UFOPA
Bolsista Programa Integrado de Ensino, Pesquisa e Extensão – PEEX/Ufopa

Paulo Roberto Brasil Santos
Doutor em Ciências Ambientais – UFOPA
Docente da UFOPA – Campus Monte Alegre, em Engenharia de Aquicultura

Suzete Roberta da Silva
Doutora em Engenharia de Pesca - UFC
Docente da UFOPA – Campus Monte Alegre, em Engenharia de Aquicultura

Realização:

Apoio:

Ano 2023